

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2004, 02:02:11 ; Search time 31 Seconds
(without alignments)
622.842 Million cell updates/sec

Title: US-09-926-805-7
Perfect score: 1949
Sequence: 1 GSHGHSRTIDARSEQDLL.....KGLEASTPSVLSQPLQLFLD 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCFUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1380	70.8	459	1	US-08-087-732-2
2	1380	70.8	459	6	Sequence 2, Appli Patent No. 5210189
3	1361	69.8	368	3	Sequence 4, Appli
4	1360	69.8	368	3	Sequence 6, Appli
5	1358	69.7	368	3	Sequence 10, Appli
6	1316	67.5	368	3	Sequence 2, Appli
7	1305	67.0	368	3	Sequence 8, Appli
8	536	27.5	945	4	Sequence 1030, Ap
9	103.5	5.3	336	4	Sequence 5549, Ap
10	101.5	5.2	2318	3	Sequence 24, Appl
11	101.5	5.2	2318	4	Sequence 24, Appl
12	97	5.0	426	4	Sequence 4401, Ap
13	96.5	5.0	575	4	Sequence 7721, Ap
14	95	4.9	1529	4	Sequence 396, App
15	93.5	4.8	453	4	Sequence 31268, A
16	92	4.7	1525	3	Sequence 2, Appli
17	92	4.7	1525	3	Sequence 2, Appli
18	92	4.7	1525	3	Sequence 2, Appli
19	91.5	4.7	684	4	Sequence 233, App
20	88.5	4.5	419	4	Sequence 17834, A
21	88	4.5	608	4	Sequence 3395, Ap
22	88	4.5	705	4	Sequence 7436, Ap
23	87.5	4.5	910	4	Sequence 2, Appli
24	87.5	4.5	1032	4	Sequence 214, App
25	87.5	4.5	2254	2	Sequence 3, Appli
26	87.5	4.5	2254	2	Sequence 3, Appli
27	87	4.5	389	4	Sequence 6456, Ap

Sequence 2, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 62, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 3693, Ap
Sequence 27566, A
Sequence 12, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 54, Appli
Sequence 53, Appli
Sequence 5110, Ap
Sequence 6678, Ap
Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-08-087-732-2
; Sequence 2, Application US/08087732
; Patent No. 5516667
; GENERAL INFORMATION:
; APPLICANT: Nishizawa, Osamu
; TITLE OF INVENTION: CHILLING RESISTANT PLANTS AND THEIR
; TITLE OF INVENTION: PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,732
; FILING DATE: 16-JUL-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP92/00024
; FILING DATE: 14-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 15883/1991
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 283807/1991
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-732-2

Query Match 70.8%; Score 1380; DB 1; Length 459;
Best Local Similarity 69.2%; Pred. No. 2.8e-140;


```

Db      241 VTGEWYPAFDASSVDNMRLIQHSVDPGHLFPLALLCHDIMPSPQVEIEGKRVIAF 300
QY      303 HGTGISVAPEINQEVTVTSCGSPPEAKAAYSQALYDSVCQYKVLHSAVHGKGLRSTP 362
Db      301 NGAGLSVAPEISPEETAATHKNPEEVREAYSKALFDSVAMQYNVLTATSKQGLGASTA 360
QY      363 SVLSQP 369
Db      361 DVLSQP 367

RESULT 4
US-09-000-092-6
; Sequence 6, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,092
; FILING DATE: 26-JAN-1998
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; PRIORITY APPLICATION NUMBER: WO PCT/JP96/01844
; FILING DATE: 03-JUL-1996
; APPLICATION NUMBER: JP 192123/1995
; FILING DATE: 27-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-000-092-6

Query Match      69.8%; Score 1360; DB 3; Length 369;
Best Local Similarity 68.4%; Pred. No. 2.8e-138;
Matches 250; Conservative 64; Mismatches 53; Indels 0; Gaps 0;

QY      3 MHGHSRTFIDARSEQDLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLSQAAPHAEDI 62
Db      1 MASHSRKFLDVRSEELLSCKKETEAGKLPNNVAGMEELYQNYRNVIESGNPKADEI 60
QY      63 VLSNMRVAFRMFLDVKEPEFSPYHEAILEPNNYMFQNYIRPLVNFRESYVGNVSFV 122
Db      61 VLSNMTVALDRILLDDPDPFVSSHKAIREPEFYIFGQNYIRPLIDFNGSVFGLSLF 120
QY      123 GVMEEQLKQDGKVVLSNHQTEADPAVIALMLETTNPHISENLIYVAGDRVITDPLCKPF 182

```

```

Db      121 KQTEKLEKQGDNIILMSNHQSEADPAVIALLEKTNLSIAENLIYTAGDRVITDPLCKPF 180
QY      183 SMGRNLLCVYSKHHNDVPELAEMKCRNTRSLKEMALLLRGSKTIWIAPSGGRDRPDP 242
Db      181 SMGRNLLCVYSKHHNDVPELAEMKCRNTRSLKEMALLLRGSKTIWIAPSGGRDRPDP 240
QY      243 ITNWFPAFPDATSLDNMRRLVDHAGLVGHYPLATLCHDIMPPLQVKEIEGKRLISF 302
Db      241 STGEWYPAFDASSVDNMRLIQHSVDPGHLFPLALLCHDIMPSPQVEIEGKRVIAF 300
QY      303 HGTGISVAPEINQEVTVTSCGSPPEAKAAYSQALYDSVCQYKVLHSAVHGKGLRSTP 362
Db      301 NGAGLSVAPEISPEETAATHKNPEEVREAYSKALFDSVAMQYNVLTATSKQGLGASTA 360
QY      363 SVLSQP 369
Db      361 DVLSQP 367

RESULT 5
US-09-000-092-10
; Sequence 10, Application US/09000092
; Patent No. 6160203
; GENERAL INFORMATION:
; APPLICANT: FERRI, Stefano
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: DNA STRANDS CODING FOR
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,092
; FILING DATE: 26-JAN-1998
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; PRIORITY APPLICATION NUMBER: WO PCT/JP96/01844
; FILING DATE: 03-JUL-1996
; APPLICATION NUMBER: JP 192123/1995
; FILING DATE: 27-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-000-092-10

Query Match      69.7%; Score 1358; DB 3; Length 368;
Best Local Similarity 68.4%; Pred. No. 4.6e-138;
Matches 249; Conservative 64; Mismatches 51; Indels 0; Gaps 0;

QY      6 HSRFTFDARSEQDLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLSQAAPHAEDI 65
Db      4 HSRKFLDVRSEELLSCKKETEAGKLPNNVAGMEELYQNYRNVIESGNPKADEI 63

```

QY 66 NMRVAFDRMFLDVKEPFEPSYHEAILEPENNYMFGNYIRPLVNFRESYVGNVSFGVM 125
 Db 64 NMTVALDRILLDVEDPFFVSSHKAIREPDYIFGQYIRPLDFGNSVGNLSFKDI 123
 QY 126 BEQKQGDQKVVILSNHQTADPAVIALMLETNPHISENIYVAGDRVITDPLCKPFSMG 185
 Db 124 EEKQKQDNIIILMNSHQSEADPAVIALLEKTNLSIAENLIYIAGDRVITDPLCKPFSMG 183
 QY 186 RNLVCVYSKKHMDVPELAEMKESNTRSLKEMALLRGSKIIWIAPSGGRDRPDITN 245
 Db 184 RNLVCVYSKKHMYDDPELVVKKRANTRSKELVLLRGSKIIWIAPSGGRDRPDVATG 243
 QY 246 QWFPAPFADATSLDNMRRLVDHAGLVGHYIPLAILCHDIMPPLQVEXEIGEKKLISPHGT 305
 Db 244 EWPAPFADASSVDNMRLIQHSDVPGLHPLPAILCHDIMPPLQVEXEIGEKKVIAFNGA 303
 QY 306 GISVAPINQEVTVASCGSPPEEAKAAYSOALYDSVCEQYKVLHSAVHGKGLFASPTSVS 365
 Db 304 GLSVAPISFEETIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKQGLGASTADVS 363
 QY 366 LSQP 369
 Db 364 LSQP 367

RESULT 6

US-09-000-092-2
 ; Sequence 2, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 368 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-000-092-2

Query Match 67.5%; Score 1316; DB 3; Length 368;
 Best Local Similarity 86.2%; Pred. No. 1.6e-133;
 Matches 243; Conservative 62; Mismatches 62; Indels 0; Gaps 0;
 QY 3 MEHGSRTFIDARSEQDLSGIQRELEAGTLPKHIAQAMBELYQNYKNVAVIQAAPHAEDI 62
 Db 1 MASHSRTYRNVRSABELISEIKRESIGRLPKSVAYAMEGLFHYRYNAVLSGGISHADEI 60
 QY 63 VLSNMRVAFDRMFLDVKEPFEPSYHEAILEPENNYMFGNYIRPLVNFRESYVGNVSF 122
 Db 61 VLSNMSVLDVLLDIEDPFPFPHKAIAREPADYISFGQYIRPLVDFGNSVGNIAIF 120
 QY 123 GVMEEOLKQGDQKVVILSNHQTADPAVIALMLETNPHISENIYVAGDRVITDPLCKPF 182
 Db 121 QEMEELKQGDNIILMNSHQSEADPAVIALLEKTNLSIAENLIYIAGDRVITDPLCKPF 180
 QY 183 SMGRNLLCVYSKKHMDVPELAEMKESNTRSLKEMALLRGSKIIWIAPSGGRDRPD 242
 Db 181 SMGRNLLCVYSKKHMYDDPELVVKKRANTRSKELVLLRGSKIIWIAPSGGRDRPDA 240
 QY 243 ITNOMEPAPFADATSLDNMRRLVDHAGLVGHYIPLAILCHDIMPPLQVEXEIGEKKLISF 302
 Db 241 VTGEWYAPFADASSVDNMRLIQHSDVPGLHPLPAILCHDIMPPLQVEXEIGEKKVIAF 300
 QY 303 HGTGISVAPINQEVTVASCGSPPEEAKAAYSOALYDSVCEQYKVLHSAVHGKGLFASPT 362
 Db 301 NGAGLSVAPISFEETIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKQGLGASTA 360
 QY 363 SVLSQP 369
 Db 361 DVLSQP 367

RESULT 7
 US-09-000-092-8
 ; Sequence 8, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 368 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-000-092-8

QY 151 ALMLETNPHISENIITYVAGDRVITDPLCKPFSMGRNLLCVYSKGMNDVPPELAEMKRS 210

US-09-091-219-24
 / Sequence 24, Application US/09091219
 / Patent No. 6171592
 / GENERAL INFORMATION:
 / APPLICANT: STUDDERT, Michael J.
 / APPLICANT: CRABB, Brendan S.
 / APPLICANT: FENG, Li
 / TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
 / FILE REFERENCE: 040268/0151
 / CURRENT APPLICATION NUMBER: US/09/091,219
 / CURRENT FILING DATE: 1998-10-05
 / EARLIER APPLICATION NUMBER: PCT/AU96/00815
 / EARLIER FILING DATE: 1996-12-18
 / EARLIER APPLICATION NUMBER: AU PN7201

; EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 US-09-091-219-24

Query Match 5.2%; Score 101.5; DB 3; Length 2318;
 Best Local Similarity 21.4%; Pred. No. 0.74; 137; Indels 103; Gaps 19;
 Matches 79; Conservative 50; Mismatches 137; Indels 103; Gaps 19;
 QY 12 DARSEQDLGSGIORELE-----AGTLPKHI-----AQAMEELYQNYKN-----49
 DB 857 ECRYNRAVNPLRGDLQVLAQKAVARTLPTSFNYGAIKATRTVELLYMKAEATYCPRPILL 916
 QY 50 AVLOSAAPHAEIDIVLS-NMRVAFDRMPL--DVKE---PFEFSPYH---EAILFEFNYYM- 99
 DB 917 AIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGFFSDFVRSNFSKLIVETINQMOE 976
 QY 100 -----FGONYIRPLVNFRESYGVNVSFGVMEQLKQGDVVLISNQHTEA-----DP 147
 DB 977 DMSTKHGPDENLVSAFEELAIGVKARTGLDEAKPWYKLIKLSRLSCMAVAARSKDP 1036
 QY 148 AVIALMETTNPHISENIYVAGDRVITDPL-----CKPFSMGRNLLCV---YSKXH 196
 DB 1037 VLVAIMLADTGLEILDSTFV---KKISDSLSLHFHPAPVPSFGAPVLLAGLVKVASSF 1093
 QY 197 MNDVPE-LAEMKKESNTRSLKEMALLRGSKII-----WIAP-----233
 DB 1094 FRSTPELEERAKQKARDINDIFAILKNGEWLVKILAIRDWIKAWIASSEKFFVTMTDL 1153
 QY 234 -----SGGRDRPDPI-----TNQWPPAPFDATSLDNMRRLVDHAGLVGHYIPLAILCHDIM 284
 DB 1154 VPGILEKQDRLNDPSKYKEKEW-----LDNARQACLKSGNV-HI---ANLCKVVA 1200
 QY 285 PPPLQVEKE 293
 DB 1201 PAPSKSRPE 1209

RESULT 11
 US-09-660-541-24
 ; Sequence 24, Application US/09660541
 ; Patent No. 6531136
 ; GENERAL INFORMATION:
 ; APPLICANT: STUDDERT, Michael J.
 ; APPLICANT: CRABB, Brendan S.
 ; APPLICANT: FENG, Li
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
 ; FILE REFERENCE: 040268/0151
 ; CURRENT APPLICATION NUMBER: US/09/660,541
 ; CURRENT FILING DATE: 2000-09-12
 ; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/091,219
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 2318
 ; TYPE: PRT
 ; ORGANISM: Foot-and-mouth disease virus
 US-09-660-541-24

Query Match 5.2%; Score 101.5; DB 4; Length 2318;
 Best Local Similarity 21.4%; Pred. No. 0.74; 137; Indels 103; Gaps 19;
 Matches 79; Conservative 50; Mismatches 137; Indels 103; Gaps 19;
 QY 12 DARSEQDLGSGIORELE-----AGTLPKHI-----AQAMEELYQNYKN-----49
 DB 857 ECRYNRAVNPLRGDLQVLAQKAVARTLPTSFNYGAIKATRTVELLYMKAEATYCPRPILL 916

QY 50 AVLOSAAPHAEIDIVLS-NMRVAFDRMPL--DVKE---PFEFSPYH---EAILFEFNYYM- 99
 DB 917 AIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGFFSDFVRSNFSKLIVETINQMOE 976
 QY 100 -----FGONYIRPLVNFRESYGVNVSFGVMEQLKQGDVVLISNQHTEA-----DP 147
 DB 977 DMSTKHGPDENLVSAFEELAIGVKARTGLDEAKPWYKLIKLSRLSCMAVAARSKDP 1036
 QY 148 AVIALMETTNPHISENIYVAGDRVITDPL-----CKPFSMGRNLLCV---YSKXH 196
 DB 1037 VLVAIMLADTGLEILDSTFV---KKISDSLSLHFHPAPVPSFGAPVLLAGLVKVASSF 1093
 QY 197 MNDVPE-LAEMKKESNTRSLKEMALLRGSKII-----WIAP-----233
 DB 1094 FRSTPELEERAKQKARDINDIFAILKNGEWLVKILAIRDWIKAWIASSEKFFVTMTDL 1153
 QY 234 -----SGGRDRPDPI-----TNQWPPAPFDATSLDNMRRLVDHAGLVGHYIPLAILCHDIM 284
 DB 1154 VPGILEKQDRLNDPSKYKEKEW-----LDNARQACLKSGNV-HI---ANLCKVVA 1200
 QY 285 PPPLQVEKE 293
 DB 1201 PAPSKSRPE 1209

RESULT 12
 US-09-134-000C-4401
 ; Sequence 4401, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR FILING DATE: 1997-08-15
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4401
 ; LENGTH: 426
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-4401

Query Match 5.0%; Score 97; DB 4; Length 426;
 Best Local Similarity 19.4%; Pred. No. 0.13;
 Matches 76; Conservative 57; Mismatches 123; Indels 136; Gaps 17;
 QY 26 ELEAGTLPKHIAQAMEELYQNYKNAVLSQSAAPHAEDIVLSNMRV---AFDR-MFLDV-- 78
 DB 96 EKAGKSLAPVAERSELVGPVTFVETRGKELEDAV--NNMKVGDVLVFENTREFVDG 154
 QY 79 -KE---PFEFSPYHEAILEFPNYMFGONYIRPLVNFRESYGVNVSFGV-MVEEQL-----129
 DB 155 KKEGNDALGKYWASLGDFVNDAGFATRAHRAHASNVIAGTGIPTVAGFLMEIKIFIG 214
 QY 130 -----KQGDVVLISNQHTEADPAVIA-----LMLE 155
 DB 215 EAVENPKRPFVAILGGAKVSDKIAVIENTLIEKADKILIGGGMAYTFMKAQGYSVGLSLE 274
 QY 156 TTNPHISENIYVAGDRVITDPLCKPFSMGRNLLCVSKGMNDVP-----ELAEWK 208
 DB 275 EKVDLAKSLMEKAGDKVLVLP-----VDTVSKFSDNAPFHTVPTSTEIPDDEE 323
 QY 209 RSN--TRSLKEMALLRGSKIIWIAPSGGRDRPDPIITNQWPPAPFDATSLDNMRRLVDH 266
 DB 324 GLDIGEKTIELFANELQAKTVMWNGPMG-----VFEMSFAK-----361
 QY 267 AGLVGHYIPLAILCHDIMPPPLQVEKEIGERKLISFHGTGISVAPEINQFVTCGSPPE 326

```
Db 362 -GTIG-----VCEAI-----ANLEDAITIIIGGD 384
QY 327 EAKAAYSQALYDSYCEQYKVLHSAVHGKGLG 358
Db 385 SAAAI-QLGVN-----KFSHISTGGASLE 410

RESULT 13
US-09-328-352-7721
; Sequence 7721, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: US/09/328,352
; NUMBER OF SEQ ID NOS: 1999-06-04
; SEQ ID NO 7721
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7721

Query Match 5.0%; Score 96.5; DB 4; Length 575;
Best Local Similarity 20.6%; Pred. No. 0.25;
Matches 68; Conservative 52; Mismatches 109; Indels 101; Gaps 14;

QY 112 RESYGVNSV-----FGVWERQLKQGDVVLISNHTQADPAVIALMLETT- 157
Db 202 QESINGNAVVKFAGESQERFYKSEBENLKKLKWIVQNLNPPVQVWACAMALIV 261
QY 158 -----NPHISENIIVAGDRVITDPLCKPFS-----MGRNLLCVYKHKHMDVPE 202
Db 262 WLALRPQILGNTTAGFVAVITAAGLSKPV-KNLTVDNEKIQGLAAHVSFELLDLPE 320
QY 203 ---LAEMKPR-----SNTSLKEMALLRGSKKIWIAPSGGRDRPDPT 244
Db 321 EONGSELKQLOGAIRFDHVLVNYADGTQAIKDFSLDIRPGETVALVGRSGA----- 372
QY 245 NQWFPAPPDATSLDNMRLVDHAGLV-----GHIYPLAILCHDIMPPLQVKEIGEKG 297
Db 373 -----KTSLVNM-----LVRFQEVSGGQIYDLDPDIRDLSSLTQTAMVNO 416
QY 298 RLISHFGTGISVAPENFOVTASCSPEE-----AKAAYSQALYDSVCEQYKVLHSAVHG 353
Db 417 QVVLFNRT---VRENIAYGQLHNA--SDSDVIAAKAAVAHDFINLPGY----DTVLG 467
QY 354 GKGLAS-----TPSVLSQPLQFLD 374
Db 469 AGLNLGGQRORAIARAILKNAPILILD 497

RESULT 14
US-09-312-283C-396
; Sequence 396, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: and Methods for Their Use
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 396
```

```
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-396
```

```
Query Match 4.9%; Score 95; DB 4; Length 1529;
Best Local Similarity 24.0%; Pred. No. 1.9;
Matches 99; Conservative 54; Mismatches 143; Indels 116; Gaps 26;
```

```
QY 32 LPKHIAQAMEEL-YQYKNVAVIQAASAA-----PHAEDIVLSNMKRVAFDRMFLDYKE-PFE- 83
Db 527 IPDHIPQYTAELRLNNEPTVLEATGIFKKLPQLRKINLSNNKIT-----DIEEGAFEG 590
QY 84 FSPYHEAI-----LEFPNYMF-QQNYIRPLV--NFRSYGVNVSVFGVMEQOLKQGDV 135
Db 581 ASGVNEILTSNRLENVQHKMEKGLSESLKTLMLRSNRISCVGNDSTFTGLGSVRL-----L 635
QY 136 VLISNHTQTEADPAVIALM--LET-----TNP-----HISENIIVAGDRVIT-DPLC-KPF 182
Db 636 SLVDNQITTVAPGAFGLTSLSLTLNLANPFCNCHLAWLGEMLRKRKRVITGNPRCKPY 695
QY 183 SMG-----RNLLCVYKHKHMDVPELAEMKRSNTRSLKEMALL---LRGSKIIW 230
Db 696 FLKEIPIQDVAIQDFTC-----DDGNDNSCSPL-----SRCESECTCLDTVVVRCNKG 746
QY 231 IAPSGGRDRPDPT-----NOMFPAPPDAT-----SLDNMRL 263
Db 747 VLPKG---IPRVTEIYLDGNQFTLVPKELSNYKHLTLIDLNNRISTLSNOSFSNMTOL 803
QY 264 VDHAGLVGHYPLAILCHD---IMPPPLQVKEIGEKEKRLISPHGTGISVAPENFOEVT 320
Db 804 L-----TLISYNRLCIPP--RTFGLKSLRLLSLHGNDISVVVPGAFCDLSA 850
QY 321 SCGSPEEAKAY-----SQALYDSVCEQYKVLHSAVHGKG-----LEASTPS 363
Db 851 LSHLAGANPLYCDNMQLSDWVKSEYKEPGIARCAPGEMADKILLTTTPS 902
```

RESULT 15

```
US-09-252-991A-31268
; Sequence 31268, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31268
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31268
```

```
Query Match 4.8%; Score 93.5; DB 4; Length 453;
Best Local Similarity 19.5%; Pred. No. 0.35;
Matches 70; Conservative 59; Mismatches 131; Indels 99; Gaps 18;
```

```
QY 5 GHSRTFFDARSEQDILLSGIQRELEAGTLPKHIAQAMEELQYKNVAVIQAASAP----- 57
Db 127 GKSVRILEASDADELFSVILEEYAKRLLENK--KKLQEIWEKTPMVMYSEAPVCAKKEE 184
QY 58 ----HAEDIVLSNMKRVAFDRMFLDVKPFEPFSPVHEA-----LLE 93
Db 185 PVQSASQDVANGUKVAVDAIAIEHAQQNAASEYGEARRYLOVDFDNGSGSTDAVVMTLE 244
QY 94 PFNYMFQNYIRPLVNFRESYVG--NVSVFQVMEQOLKQGDV---VLISNHTQTEAD-- 146
```

Db	245	PTNGGYNQYVLIRNGDAWEAKTRLDIMNGASDLADGEGVLSVIELSHGPDADCC	304
QY	147	PAVIALMETTNPHISEN-----IIVAGRVITPLCKPFSMGRNLLCVTSKXHMNDV	200
Db	305	PSMETKRLYKNGSVLQFYKNDGAVPTQASSKSLAP--PF-----YSVLTLND	352
QY	201	PELAEMKKRSNTRSLKEWALLLRGSKLIW-----IAPSGGRD-----RPDP	242
Db	353	SD-----EKVTTKRIELCSM-----ASKVGQLDSIDIILAKINRTGLANINPFI	404
QY	243	---ITNQWFPAPFDATSLDMRNLVDHAGLVGHYFLAILCHDI--WPPLQVREIGE	296
Db	405	YHQMICKDW-----SLKDMNRASQDEYLVG-IYN-SFECHKLHQEPAIQVSDVPG	453

Search completed: March 16, 2004, 02:08:14
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2004, 00:53:55 ; Search time 26 Seconds
(without alignment)
749.009 Million cell updates/sec

Title: US-09-926-805-7

Perfect score: 1949

Sequence: 1 GSMHGHSTFIDARSEQD...KGLEASTPSVLSQPLQLD 374

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1913.5	98.2	463	1	PLSB CARTI
2	1389	71.3	459	1	PLSB ARATH
3	1379	70.8	457	1	PLSB PEA
4	1377	70.7	470	1	PLSB CUOSA
5	1350	69.3	396	1	PLSB CUCMO
6	1332	68.3	461	1	PLSB PHAVU
7	1321	67.8	472	1	PLSB SPIOL
8	103.5	5.3	882	1	PPDK RICCN
9	101.5	5.2	539	1	CH60 BACTR
10	101.5	5.2	2332	1	POLG FMDVO
11	98	5.0	2332	1	POLG FMDVA
12	97.5	5.0	545	1	CH60 OCEIH
13	96	4.9	432	1	SYH PYRFU
14	95.5	4.9	880	1	PODK RICPR
15	94	4.8	1061	1	EC4A ARATH
16	93.5	4.8	539	1	CH60 BACST
17	93.5	4.8	604	1	SP20 YEAST
18	92.5	4.7	400	1	PGK LACPL
19	92.5	4.7	542	1	CH60 LISIN
20	92.5	4.7	2748	1	NUM1 YEAST
21	92	4.7	1529	1	SLT2 HUMAN
22	92	4.7	1571	1	ATG5 YEAST
23	91.5	4.7	448	1	TX19 HUMAN
24	91.5	4.7	543	1	CH60 CIOAB
25	91.5	4.7	684	1	CPSC HUMAN
26	91.5	4.7	810	1	SC12 ARATH
27	91.5	4.7	950	1	ATC1 YEAST
28	91	4.7	219	1	P1MT PYRAB
29	91	4.7	543	1	CH60 BACSU
30	91	4.7	1061	1	EC4A ARATH
31	91	4.7	2333	1	POLG FMDV1
32	91	4.7	2336	1	POLG FMDVZ
33	90.5	4.6	394	1	ALG2_PSEFL

34 90.5 4.6 542 1 CH60 LISMO
35 90.5 4.6 1376 1 RPOD ARATH
36 90 4.6 544 1 CH60 BACHD
37 89.5 4.6 684 1 CPSC BOVIN
38 89 4.6 783 1 NQO3 THERM
39 89 4.6 1690 1 RPOC THERM
40 89 4.6 2128 1 SPCB MOUSE
41 89 4.6 2183 1 DO10 HUMAN
42 88.5 4.5 332 1 LIPB KUULA
43 88.5 4.5 427 1 MURA BORBU
44 88.5 4.5 431 1 SYH PYRAB
45 88 4.5 1384 1 RPOD SINAI

ALIGNMENTS

RESULT 1

PLSB CARTI STANDARD; PRT; 463 AA.
AC Q42713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycero1-3-phosphate acyltransferase, chloroplast precursor
DE (EC 2.3.1.15) (GPAT)
OS Carthamus tinctorius (Safflower)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Carthamus.
OX NCBI_TaxID=4222;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=95148760; PubMed=7846182;
RA Shella R.S., Mackenzie S.L.;
RT "Nucleotide sequence of a cDNA from Carthamus tinctorius encoding a
RL Glycero1-3-phosphate acyl transferase.";
CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
of glycerol-3-phosphate. The enzyme from chilling-resistant plants
discriminate against non-fluid palmitic acid and select oleic acid
whereas the enzyme from sensitive plants accepts both fatty acids.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
also function in the regulation of membrane biogenesis.
CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -!- SIMILARITY: Belongs to the plant GPAT family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L33841; AAA74319.1;
InterPro: IPR002123; Acyltransferase.
Pfam: PF01553; Acyltransferase; 1.
SMART: SM00563; PlsC; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
Transit peptide; Chloroplast.
FT TRANSIT 1 91 CHLOROPLAST (POTENTIAL).
FT CHAIN 92 463 GLYCERO1-3-PHOSPHATE ACYLTRANSFERASE.
SQ SEQUENCE 463 AA; 50834 MW; 5792E933068A534D CRC64;

Query Match 98.2%; Score 1913.5; DB 1; Length 463;
Best Local Similarity 98.7%; Fred. No. 1.5e-150;
Matches 370; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GSMHGHSTFIDARSEQD...KGLEASTPSVLSQPLQLD 374

Db 89 GSDHGRSTFDARSEQDLSGIORELEAGTLPKHIAQAMEELYQNVKAVLQSAAPHAE 148
 QY 61 DIVLSNMRVAFDRFLDKPEFSPYHEALLESFNYMFGQNYRPLVNPRESYGVNS 120
 Db 149 DIVLSNMRVAFDRFLDKPEFSPYHEALLESFNYMFGQNYRPLVNPRESYGVNS 208
 QY 121 VFGWEEQLQKQDKVLSNQHTEADPAVIALMETTNPHISENIIVAGDRVITDPLCK 180
 Db 209 VFGWEEQLQKQDKVLSNQHTEADPAVIALMETTNPHISENIIVAGDRVITDPLCK 268
 QY 181 PFSMGRNLLCVYKXKNDVDELAKMKRSTRSLK-EMALLRGSKIIWIAPSGGRDR 239
 Db 269 PFSMGRNLLCVYKXKNDVDELAKMKRSTRSLKGRMALLRGSKIIWIAPSGGRDR 328
 QY 240 PDPITNWFPAFPDTSLDNNRRLVDHAGLVGHLYPLAILCHDIMPPLQVEKEIGKRL 299
 Db 329 PDPITNWFPAFPDTSLDNNRRLVDHAGLVGHLYPLAILCHDIMPPLQVEKEIGKSW 388
 QY 300 ISPHGTGISVAPEINQEVNTASCSPPEAKAAYSQALYDSYCEQYKVLHSAVHGKGLEA 359
 Db 389 ISPHGTGISVAPEINQEVNTASCSPPEAKAAYSQALYDSYCEQYKVLHSAVHGKGLEA 448
 QY 360 STPSVLSQPLQFLD 374
 Db 449 STPSVLSQPLQFLD 463

RESULT 2

PLSB ARATH STANDARD; PRT; 459 AA.
 AC Q43307; G3FVR5;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 GN AT51 OR AT1G32200 OR F3C3.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93144702; PubMed=7678766;
 RA Nishida I., Tasaka Y., Shiraiishi H., Murata N.;
 RT "The gene and the RNA for the precursor to the plastid-located
 RT glycerol-3-phosphate acyltransferase of Arabidopsis thaliana";
 RL Plant Mol. Biol. 21:267-277(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana";
 RL Nature 408:816-820(2000).

[3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Cheuk R.F.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Yu G.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallenberg E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome";
 RL Science 302:842-846(2003).
 CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 CC discriminates against non-fluid palmitic acid and select oleic acid
 CC whereas the enzyme from sensitive plants accepts both fatty acids.
 CC This is an oleate-selective acyltransferase.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 CC acyl-sn-glycerol 3-phosphate.
 CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 CC also function in the regulation of membrane biogenesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -1- SIMILARITY: Belongs to the plant GPAT family.
 CC -----
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00673; BAA00576.1; -;
 DR EMBL; D00672; BAA00575.1; -;
 DR EMBL; AC084165; AAG23437.1; -;
 DR EMBL; AY093169; AAM13168.1; -;
 DR EMBL; BT008758; AAP49520.1; -;
 DR PIR; E86446; E86446;
 DR PIR; S31083; S31083;
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 DR SMART; SM00563; Plac; 1.
 DR KX Phospholipid biosynthesis; Transferase; Acyltransferase;
 KW Transit peptide; Chloroplast.
 FT TRANSIT 1 90 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 91 459 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
 FT CONFLICT 93 93 S -> N (IN REF. 1).
 FT CONFLICT 287 287 D -> V (IN REF. 1).
 SQ SEQUENCE 459 AA; 50421 MW; 684CF97EA5B82A7E CRC64;
 Query Match 71.3%; Score 1389; DB 1; Length 459;
 Best Local Similarity 69.5%; Pred. No. 3.5e-107;
 Matches 253; Conservative 57; Mismatches 54; Indels 0; Gaps 0;
 QY 6 H8RTFDARSEQDLSGIORELEAGTLPKHIAQAMEELYQNVKAVLQSAAPHAEIVLS 65
 Db 94 H8RTFDARSEQDLSGIGKEAEAGLPAVNAAGHEELYWYKNVLSGASRADETUVS 153
 QY 66 NMRVAFDRMFLDKPEFSPYHEALLESFNYMFGQNYRPLVNPRESYGVNSVFGVM 125
 Db 154 NMSVAFDRMLLGVEDPYTFNPYKAVREPFDYMFVHTVIRPLIDFKNSYVGNASIFSEL 213
 QY 126 EQQLQKQDKVLSNQHTEADPAVIALMETTNPHISENIIVAGDRVITDPLCKPFSMG 185
 Db 214 EDKIQGHNVLLISNQHSEADPAVLSLLEAQSPFIGNIKCVAGDRVITDPLCKPFSMG 273

186 RNLVCYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPDITN 245
 274 RNLVCYKXKHMNDVDELAEKMKRSNTRSLKEMATMLRSGQLIWIAPSGRRDRPNSTG 333
 246 QWFPAPFDATSLDNMRRLVHAGLVGHYPLAILCHDIMPPLLOVEKEIGEKELISFHGT 305
 334 EWFPAFDASSVDNMRRLVHSGAPGHYPMSLCYDIMPFPQVKEIGEKELVGFHGT 393
 306 GISVAPEINFQEVATSCGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTPSVS 365
 394 GSIAPAEINFSDVTADCESPNEAKAAYSQALYKSVNEQYELNSAIKRRGVEASTSRVS 453
 366 LSQP 369
 454 LSQP 457

Query Match 70.8%; Score 1379; DB 1; Length 457;
 Best Local Similarity 71.1%; Pred. No. 2.3e-106; Indels 0; Gaps 0;
 Matches 258; Conservative 53; Mismatches 52;

QY 7 SRTFDARSEQDLSGIGRELEAGTLPHKIAQAMEELYQYKNAVLQSAAPHAEDIVLSN 66
 Db 93 SRTFLNAQNEQDVLGIGKEVEAGTLPASIAAGMEEVLYNKSAVIGKDPKANEIVLSN 152
 QY 67 MRVAFDRMFLDVKPEPESFYHEALILEPENYMFQNYIRPLVNFRESYVGVNVPVGM 126
 Db 153 MTALLDRIFLDVKEPVPFAHHKAKREPPDYMFQNYIRPLVDFETSVMGNMPLFIQME 212
 QY 127 EQLKQDKVVLISNHQTEADPAVIALMLTTPPHISENIYVAGDRVITDPLCKPFSMGR 186
 Db 213 EQLKQGHNILLNSHQSEADPAIALLLEMLPHIAENLIYVAGDRVITVPLCKPFSIGR 272
 QY 187 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPDITNQ 246
 Db 273 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPVANSCE 332
 QY 247 WPAFPDATSLLDNMRRLVHAGLVGHYPLAILCHDIMPPLLOVEKEIGEKELISFHGTG 306
 Db 333 WAPAFDSSVDNMRRLVHSGAPGHYPLAILCHDIMPPLLOVEKEIGEKELISYHGTG 392
 QY 307 ISVAPEINFQEVATSCGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTPSVSL 366
 Db 393 ISTAPEISFNTTAACENPERAKDAYTRALYDSVTEQYDVLKSAIHGKGLQASTPVSL 452
 QY 367 SQP 369
 Db 453 SQP 455

RESULT 4
 PLSB CUCA
 ID PLSB CUCA STANDARD; PRT; 470 AA.
 AC Q39639;
 DT 15-DEC-1998 (Rel. 37; Created)
 DT 15-DEC-1998 (Rel. 37; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID:3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson T.C., Schneider J.C., Somerville C.;
 RT "Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-
 phosphate acyltransferase from cucumber.";
 RL Plant Physiol. 99:771-772(1992).
 CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 discriminate against non-fluid palmitic acid and select oleic acid
 whereas the enzyme from sensitive plants accepts both fatty acids.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 acyl-sn-glycerol 3-phosphate.
 CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 also function in the regulation of membrane biogenesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: Belongs to the plant GPAT family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC
 EMBL; X59041; CAA41769.1; -
 DR F0182339; S18239.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.

186 RNLVCYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPDITN 245
 274 RNLVCYKXKHMNDVDELAEKMKRSNTRSLKEMATMLRSGQLIWIAPSGRRDRPNSTG 333
 246 QWFPAPFDATSLDNMRRLVHAGLVGHYPLAILCHDIMPPLLOVEKEIGEKELISFHGT 305
 334 EWFPAFDASSVDNMRRLVHSGAPGHYPMSLCYDIMPFPQVKEIGEKELVGFHGT 393
 306 GISVAPEINFQEVATSCGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTPSVS 365
 394 GSIAPAEINFSDVTADCESPNEAKAAYSQALYKSVNEQYELNSAIKRRGVEASTSRVS 453
 366 LSQP 369
 454 LSQP 457

Query Match 70.8%; Score 1379; DB 1; Length 457;
 Best Local Similarity 71.1%; Pred. No. 2.3e-106; Indels 0; Gaps 0;
 Matches 258; Conservative 53; Mismatches 52;

QY 7 SRTFDARSEQDLSGIGRELEAGTLPHKIAQAMEELYQYKNAVLQSAAPHAEDIVLSN 66
 Db 93 SRTFLNAQNEQDVLGIGKEVEAGTLPASIAAGMEEVLYNKSAVIGKDPKANEIVLSN 152
 QY 67 MRVAFDRMFLDVKPEPESFYHEALILEPENYMFQNYIRPLVNFRESYVGVNVPVGM 126
 Db 153 MTALLDRIFLDVKEPVPFAHHKAKREPPDYMFQNYIRPLVDFETSVMGNMPLFIQME 212
 QY 127 EQLKQDKVVLISNHQTEADPAVIALMLTTPPHISENIYVAGDRVITDPLCKPFSMGR 186
 Db 213 EQLKQGHNILLNSHQSEADPAIALLLEMLPHIAENLIYVAGDRVITVPLCKPFSIGR 272
 QY 187 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPDITNQ 246
 Db 273 NLLCVYKXKHMNDVDELAEKMKRSNTRSLKEMALLRGGSKIWIAPSGRRDRPVANSCE 332
 QY 247 WPAFPDATSLLDNMRRLVHAGLVGHYPLAILCHDIMPPLLOVEKEIGEKELISFHGTG 306
 Db 333 WAPAFDSSVDNMRRLVHSGAPGHYPLAILCHDIMPPLLOVEKEIGEKELISYHGTG 392
 QY 307 ISVAPEINFQEVATSCGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLFASTPSVSL 366
 Db 393 ISTAPEISFNTTAACENPERAKDAYTRALYDSVTEQYDVLKSAIHGKGLQASTPVSL 452
 QY 367 SQP 369
 Db 453 SQP 455

RESULT 4
 PLSB CUCA
 ID PLSB CUCA STANDARD; PRT; 470 AA.
 AC Q39639;
 DT 15-DEC-1998 (Rel. 37; Created)
 DT 15-DEC-1998 (Rel. 37; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID:3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson T.C., Schneider J.C., Somerville C.;
 RT "Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-
 phosphate acyltransferase from cucumber.";
 RL Plant Physiol. 99:771-772(1992).
 CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 discriminate against non-fluid palmitic acid and select oleic acid
 whereas the enzyme from sensitive plants accepts both fatty acids.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 acyl-sn-glycerol 3-phosphate.
 CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 also function in the regulation of membrane biogenesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: Belongs to the plant GPAT family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC
 EMBL; X59041; CAA41769.1; -
 DR F0182339; S18239.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M80571; AAA33122.1; -
 DR PIR; T10193; T10193.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 DR SMART; SM00563; Pfam; 1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 FT Transit peptide; Chloroplast.
 FT TRANSIT 1 101 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 102 470 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
 SQ SEQUENCE 470 AA; 51875 MW; AF60144E19D2B39B CRC64;

Query Match 70.7%; Score 1377; DB 1; Length 470;
 Best Local Similarity 69.5%; Pred. No. 3.5e-106;
 Matches 253; Conservative 59; Mismatches 52; Indels 0; Gaps 0;

QY	6	HSRTFIDARSDLLSGIORELEAGTLPKHIAQAMEELYQYKNVLSQAAPHAEIVLS	65
DB	105	HSRAFLDLRSEELLSCIRRETEAGKLPNSVAAGMEELYQYKNVPSGNPKADEIVLS	164
QY	66	NMRVAFDRMFLDVKEPEFSPYHEALIEPFFNYMFGQYIRPLNFRRESYVGVNVSFVGM	125
DB	165	NMTVALDRILLDVEDPFFVSSHKAIREPFDYITFGQYVRLDFDENSEFVGNLSLFKDI	224
QY	126	EEQLKQGDQKVLISNHQTEADPAVIALMETTNPHISENIIVYAGDRVITDPLCKPFSMG	185
DB	225	EEKLHQGHNVLLISNHQTEADPAISLLEKTNPYIAENMIVYAGDRVIADPLCKPFSIG	284
QY	186	RNLICVYSKGMNDVPELAEMKKRNTSLKEMALLLRGSKIIWIAPSGGRDRPDPTN	245
DB	285	RNLICVYSKGMNDVPELAETPKRANTSLKEMALLLRGSKIIWIAPSGGRDRPDPTG	344
QY	246	QWFPAPDTSNLMRRLVHAGLVGHYIPLAILCHDIMPFPQVEKEIGKRLISFHGT	305
DB	345	EWTPAPDASSVDMRRLVHAGLVGHYIPLAILCHDIMPFPQVEKEIGKRLISFHGT	404
QY	306	GLSVAPINFEQVTSACGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLASPTSVS	365
DB	405	GLSVAPISFDEIAASRDNDPEVREAYSKALYDSVAKQYVNLKAAIDGKQLEASVADVI	464
QY	366	LSQP 369	
DB	465	LSQP 468	

RESULT 5
 ID_PLSB_CUCMO STANDARD; PRT; 396 AA.
 AC P10349;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
 DE (EC 2.3.1.15) (GPAT).
 OS Cucurbita moschata (Cushaw squash) (Winter crookneck squash).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
 OX NCBI_TaxID=3662;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. Shirakikuza; TISSUE=Cotyledon;
 RX MEDLINE=89005726; PubMed=2458971;
 RA Ishizaki O., Nishida I., Agata K., Eguchi G., Murata N.;
 RT "Cloning and nucleotide sequence of cDNA for the plastid glycerol-3-
 RT phosphate acyltransferase from squash."
 RL FEBS Lett. 238:424-430 (1988).
 RN [2]

RP CHARACTERIZATION.
 RA Frentzen M., Nishida I., Murata N.;
 RT "Properties of the plastidial acyl-(acyl-carrier-protein):
 RT glycerol-3-phosphate acyltransferase from the chilling-sensitive
 RT plant squash (Cucurbita moschata).";
 RL Plant Cell Physiol. 28:1195-1201 (1987).
 CC -!- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
 CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants
 CC discriminate against non-fluid palmitic acid and select oleic acid
 CC whereas the enzyme from sensitive plants accepts both fatty acids.
 CC Squash is chilling-sensitive.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
 CC acyl-sn-glycerol 3-phosphate.
 CC -!- PATHWAY: First step in de novo phospholipid biosynthesis. It may
 CC also function in the regulation of membrane biogenesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- SIMILARITY: Belongs to the plant GPAT family.
 CC -!- CAUTION: May represent a 5'-truncated sequence in which part of
 CC the leader sequence is missing.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Y00771; GAA69740.1; -
 DR PIR; S01660; S01660.
 DR PDB; 1K30; 31-OCT-01.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 DR SMART; SM00563; Pfam; 1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 KW Transit peptide; Chloroplast; 3D-structure.
 FT TRANSIT 1 28 CHLOROPLAST.
 FT CHAIN 29 396 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
 SQ SEQUENCE 396 AA; 43838 MW; CB96E8C057B88112 CRC64;

Query Match 69.3%; Score 1350; DB 1; Length 396;
 Best Local Similarity 67.6%; Pred. No. 4.8e-104;
 Matches 246; Conservative 67; Mismatches 51; Indels 0; Gaps 0;

QY	6	HSRTFIDARSDLLSGIORELEAGTLPKHIAQAMEELYQYKNVLSQAAPHAEIVLS	65
DB	32	HSRKFLDVASEELLSCIRRETEAGKLPNSVAAGMEELYQYKNVPSGNPKADEIVLS	91
QY	66	NMRVAFDRMFLDVKEPEFSPYHEALIEPFFNYMFGQYIRPLNFRRESYVGVNVSFVGM	125
DB	92	NMTVALDRILLDVEDPFFVSSHKAIREPFDYITFGQYVRLDFDENSEFVGNLSLFKDI	151
QY	126	EEQLKQGDQKVLISNHQTEADPAVIALMETTNPHISENIIVYAGDRVITDPLCKPFSMG	185
DB	152	EEKLHQGHNVLLISNHQTEADPAISLLEKTNPYIAENMIVYAGDRVIADPLCKPFSIG	211
QY	186	RNLICVYSKGMNDVPELAEMKKRNTSLKEMALLLRGSKIIWIAPSGGRDRPDPTN	245
DB	212	RNLICVYSKGMNDVPELAETPKRANTSLKEMALLLRGSKIIWIAPSGGRDRPDPTG	271
QY	246	QWFPAPDTSNLMRRLVHAGLVGHYIPLAILCHDIMPFPQVEKEIGKRLISFHGT	305
DB	272	EWTPAPDASSVDMRRLVHAGLVGHYIPLAILCHDIMPFPQVEKEIGKRLISFHGT	331
QY	306	GLSVAPINFEQVTSACGSPPEAKAAYSQALYDSVCEQYKVLHSAVHGKGLASPTSVS	365
DB	332	GLSVAPISFDEIAASRDNDPEVREAYSKALYDSVAKQYVNLKAAIDGKQLEASVADVI	391
QY	366	LSQP 369	
DB	392	LSQP 395	

```

RESULT 6
PLSB PHAVU
ID PLSB PHAVU STANDARD; PRT; 461 AA.
AC Q43822;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, chloroplast precursor
DE (EC 2.3.1.15) (GPAT).
GN PLSB.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OC NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Annabel; TISSUE=Leaf;
RX MEDLINE=95232196; PubMed=7716242;
RA Fritz M., Heinz E., Wolter F.P.;
RT "Cloning and sequencing of a full-length cDNA coding for
RT sn-glycerol-3-phosphate acyltransferase from Phaseolus vulgaris.";
RL Plant Physiol. 107:1039-1040(1995).
CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position
CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants
CC discriminates against non-fluid palmitic acid and select oleic acid
CC whereas the enzyme from sensitive plants accepts both fatty acids.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may
CC also function in the regulation of membrane biogenesis.
CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -1- SIMILARITY: Belongs to the plant GPAT family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79722; CAA56159.1; -
CC PIR; T11819; T11819.
CC InterPro; IPR002123; Acyltransferase.
CC Pfam; PF01553; Acyltransferase; 1.
CC SMART; SM00563; PlsC; 1.
CC Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Transit peptide; Chloroplast. CHLOROPLAST (POTENTIAL).
CC TRANSIT 1 96
CC CHAIN 97 461 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
CC SEQUENCE 461 AA; 50697 MW; EAC4FC837908B38A CRC64;
Query Match 68.3%; Score 1332; DB 1; Length 461;
Best Local Similarity 67.2%; Pred. No. 1.8e-102;
Matches 244; Conservative 61; Mismatches 58; Indels 0; Gaps 0;
QY 7 SRTFDARSEQDLGSIQRELEAGTLPKHIAQAAEELYNKYNAVYQSAAPHAEIVLSN 66
DB 97 SRTFLNAQSEQDVFAKKEVEAGSLPANVAAGMEEVYNNYKAVIQSGDPFANEIVLSN 156
QY 67 MRVAFDRFLVDKPEFSPVHEALELPFNMYFQNYIRPLNFRSIVGNVSVFGVME 126
DB 157 MIALDRVFLVDTPDFVFPKHKAKREPFDYVVFQNYIRPLVDPKNAVYGNMPLFIEME 216
QY 127 EQLKGDKGVLLISNQTADPAVIALMETTNPHSENIIVAGDRVITDPLCKPFGMR 186
DB 217 EKLKGHLLIMNSNQTADPAISLLETLPLPYTAENLTIVAGDRVITDPLSKPFSIGR 276
QY 187 NLLCVYSKHMNDVPPELAEMKRSNTRSLKEMALLRGSKIWTAPSGGRDRPDPTNQ 246
DB 277 NLCVYSKHMDDPALVEMKRTANIRALKEMALLRNGSLQVWVITAPSGGRDRPDQTR 336

```

us-09-926-805-7.rsp

Page 5

RESULT 7

PLSB SPIOL

ID PLSB SPIOL STANDARD; PRT; 472 AA.

AC Q43869;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glycerol-3-phosphate acyltransferase, chloroplast precursor

DE (EC 2.3.1.15) (GPAT).

GN GAT OR ACT1.

OS Spinacia oleracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllales; Amaranthaceae; Spinacia.

OC NCBI_TaxID=3562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Bifurca; TISSUE=Seedling;

RA Ishikazi-Nishizawa O., Azuma M., Ohtani T., Murata N., Toguri T.;

RT "Nucleotide sequence of cDNA from Spinacia oleracea encoding plastid

RT glycerol-3-phosphate acyltransferase.";

RL (in) Plant Gene Register PGR95-014.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Melody; TISSUE=Leaf;

RA Wolter F.P.;

RT "Conserved intron position in 3' untranslated region of a cDNA

RT encoding the plastidial sn-glycerol-3-phosphate acyltransferase of

RT spinach.";

RL (in) Plant Gene Register PGR96-118.

RN [3]

RP CHARACTERIZATION.

RX MEDLINE=83131613; PubMed=6825679;

RA Frentzen M., Heinz E., McKeon T.A., Stumpf P.K.;

RT "Specificities and selectivities of glycerol-3-phosphate

RT acyltransferase and monoacylglycerol-3-phosphate acyltransferase from

RT pea and spinach chloroplasts.";

RL Eur. J. Biochem. 129:629-636(1983).

CC -1- FUNCTION: Esterifies acyl-group from acyl-ACP to the sn-1 position

CC of glycerol-3-phosphate. The enzyme from chilling-resistant plants

CC discriminates against non-fluid palmitic acid and select oleic acid

CC whereas the enzyme from sensitive plants accepts both fatty acids.

CC This is an oleate-selective acyltransferase.

CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-

CC acyl-sn-glycerol 3-phosphate.

CC -1- PATHWAY: First step in de novo phospholipid biosynthesis. It may

CC also function in the regulation of membrane biogenesis.

CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.

CC -1- SIMILARITY: Belongs to the plant GPAT family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; X77370; CAA54559.1; -

CC EMBL; Z49091; CAA88913.1; -


```

QY 237 -----RD-----RPD-----PITNWFAPFAPDATSLDNMR-RLVDHAGLVGHYI 274
Db 578 NKIPLVDMIIAPDIERRKLAQVQLLPLOYEDFVALFRVMDKPVNRLDDP----- 629
QY 275 PLALCHDIMPPLQVKEIG-----EKLSFHGTGTSVAPEINFQEVTA SCGS 324
Db 630 PL-----HEFLPTTDEDKNLANSINLPLSMINQRLHAMH-----EVNPMLGHRGCLGICS 681
QY 325 PEEAKAAYS---QALYDSVCEQYKVLHSAVHGKGLASPTSPVS 365
Db 682 PE-----IYQWQIEAIFTAIFE-----LHKXSHIECNLEIMPLIS 717

RESULT 9
ID CH60 BACTR STANDARD; PRT; 539 AA.
AC QRVV84;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 60 kDa chaperonin (Protein Cpn60) (GroEL protein).
GN GROEL OR GROEL.
OS Bacillus thermoglucosidasius (Geobacillus thermoglucosidasius).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1426;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=KP1006;
RX MEDLINE=21823396; PubMed=11834128;
RA Watanabe K., Fujiwara H., Inui K., Suzuki Y.;
RT "oligo-1,6-glucosidase from a thermophile, Bacillus
thermoglucosidasius KP1006, was efficiently produced by combinatorial
expression of GroEL in Escherichia coli.";
RL Biotechnol. Appl. Biochem. 35:35-43(2002).
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and
proper assembly of unfolded polypeptides generated under stress
conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
7 subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AB025944; BAB83940.1; -
DR HAMAP; MF 00600; -; 1.
DR InterPro; IPR001844; Chaprin Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; cpn60 TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
SQ SEQUENCE 539 AA; 57192 MW; 70GBE183F07C67B5 CRC64;

Query Match 5.2%; Score 101.5; DB 1; Length 539;
Best Local Similarity 19.3%; Pred. No. 1.2;
Matches 91; Conservative 63; Mismatches 131; Indels 187; Gaps 20;

QY 22 GIQRELEAGTIPKHTAQWELIYQ-----NYKNAVLSQAAPHAED-----IVLSNMKVA 70
Db 114 GIKKGIE-----KAVAAVEELKALSKPIQKESIAQVAALSADEEVQLIAEMERVG 168
QY 71 FDRMF-----LDVKEPFE-----SPY-----HEALEPPNYMFGQNVIRP 107

```

```

Db 169 NDGVITLESKGTTELDVVVEGMOFGYASPYMITDEKQEA VLE-----NP 216
QY 108 LVNFRESYGVNV-SVFGMEEQLKQGDVKVLI SNH-QTEADPAVIALMLTETNP HIS--- 162
Db 217 YILITDKISNIQDILPILQVQVQCKPLLIADVEGEALATLVNKLGRGTTFAVAKA 276
QY 163 -----ENIIVAGDRVITDPLCKPF-----SMGRNLLCVYSKKH--- 196
Db 277 PGFGDRRKAMLEIDAILTGGEVISEELGRELKSTTIASLG RASKVVVTKENTTIVEGAGD 336
QY 197 -----MNDV-----PELAEMKKR--- 209
Db 337 SERIKARINQIRAOLETTSEFDRGKLOERLAKLAGGAVIKVGAATETELKERKRIED 396
QY 210 ---SNTSLKEMALLRGG-----SKIWIAPSGGRDRPDPTITNWFAPFAPDATSLDNM 260
Db 397 ALNSTRAAVEEGIVAGGVTALMNVNKAIEAEGDEATCVKIVLRAIEP-----V 448
QY 261 RLIVDHAGLVGHYIPLAILCHDIMPPLPQVEKEIGKRLISFHGTGTSVAPEINFQEVTA 320
Db 449 ROIQAQVAGLEGSVI-----VERLKSEK-----PGIGFNAATG 480
QY 321 SCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGKGLER---STPSVSLSQP 369
Db 481 EWNWMLER-----GIVDTKVRSAQNAAASVAAMFLTTEAVVADKP 522

RESULT 10
POLG FMDVCO
ID POLG FMDVCO STANDARD; PRT; 2332 AA.
AC P03305;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Genome polypeptide [Contains: Nonstructural protein P20A; Coat
protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;
Core protein P12; Core protein P34; Core protein P14; Genome-linked
protein VP6; Protease (EC 3.4.22.-); RNA-directed RNA polymerase
(EC 2.7.7.48)].
DE Foot-and-mouth disease virus (strain O1) (Aphthovirus O) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=73482;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=O1K;
RX MEDLINE=84297249; PubMed=6089122;
RA Fores S., Strebel K., Beck E., Schaller H.;
RT "Nucleotide sequence and genome organization of foot-and-mouth
disease virus.";
RL Nucleic Acids Res. 12:6587-6601(1984).
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=O1BFS;
RX MEDLINE=83143292; PubMed=6298715;
RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RT "Comparison of the amino acid sequence of the major immunogen from
three serotypes of foot and mouth disease virus.";
RL Nucleic Acids Res. 10:8285-8295(1982).
RN [3]
RC X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=89143740; PubMed=2537470;
RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;
RT "The three-dimensional structure of foot-and-mouth disease virus at
2.9-A resolution.";
RL Nature 337:709-716(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.

```


RESULT 12

```

CH60_OCEIH
ID CH60_OCEIH STANDARD; PRT; 545 AA.
AC Q8CKL3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60 kDa chaperonin (protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR OR0656.
OS Oceanobacillus ihewensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus ihewensis isolated from the Iheya
RT ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AP004595; BAC12612.1; -.
DR HAMAP; MF 00600; -; 1.
DR InterPro; IPR001844; Chaprln Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00298; CHAPERONING60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; FALSE_NEG.
DR Chaperone; ATP-binding; Complete proteome.
DR KW Chaperone; ATP-binding; Complete proteome.
DR SQ SEQUENCE 545 AA; 57508 MW; 311362C2C485E6AD0 CRC64;

Query Match 5.0%; Score 97.5; DB 1; Length 545;
Best Local similarity 20.6%; Pred. No. 2.6;
Matches 95; Conservative 66; Mismatches 130; Indels 171; Gaps 23;

QY 22 GIQRELEAGTLP-KHIAQAMELYQNYKNVLOSAAPHAED-----IVLSMRVAFORM 74
Db 118 GLEKAVELAQBLKGISPIES-----KEAISQIAAVSSGDEEVGQLTAEMERVNDGV 172
QY 75 F-----LDVKEPEF-----SPY-----HEAILEPFNYMFGQNYRLVNF 111
Db 173 ITIESKGFNTELEVVEGQDRGYASPYWTDQDKMEAVLED-----PYILI 220
QY 112 RESYVGNV-SVEGVNMEQLKQDGKVVLLSNH-QTEADPAVIALMLETNPHTIS----- 162
Db 221 TDKKIGNIQVLPVLEQVVQQKPLLMIAEDVEGEALATLVNKLRTGFNAVAKPCFG 280
QY 163 -----ENIIVAGDRVITDPL-----CKPFS-----MGRNLLCVYSGKMNIV-----PE-- 202
Db 281 DRKAMLEDIAVLTAERVITDGLDGLKLSAIDGLGRASKVYVTKENTYVEVSGDPEAI 340
QY 203 ----LAEMKRS-----NTRSLKEMALLRGGSKIWI----- 231
Db 341 SRVIAIQAQAEESTSDFKELQERLAKLAGGVAVIKVGAATTELKERKLRTEDALNS 400
QY 232 -----APSGG-----RDRPDPITNWFAPPAFATSL-----DNNRRLLVDHAGLVGH 272

```

```

Db 401 TRAGVEGMEVGGGTAIVNIHQVSELTEGDEATGASIVLRALSEPVRQIVHNAGLEGS 460
Qy 273 IVPLAILCHDIMPPLQVEKEIGKRLISFGHTGTSVAPE--INFOEVASCGSPPEAKA 330
Db 461 II-----VERLKGK-----VGIGYNAATDEWNNVEA-----488
Qy 331 AYSQALYSVCEQYKVLHSAVHGKGLA---STPSVSLSP 369
Db 489 -----GIVDTKVTRSALQNAASVAAMFLTTEAVVADKP 522

```

RESULT 13

```

ID SYH PYRFU STANDARD; PRT; 432 AA.
AC Q8U431;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidyl-tRNA synthetase (SC 6.1.1.21) (Histidine--tRNA ligase)
DE (HisRS).
GN HIS OR PF0264.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RX Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RA "The complete sequence of the Pyrococcus furiosus genome.";
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -! SUBCELLULAR LOCATION: Cytoplasmic.
CC -! SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AE010150; AAL80398.1; -
CC HAMAP; MF 00127; -; 1.
CC InterPro; IPR004154; HGTP anticodon.
CC InterPro; IPR004516; HisS
CC InterPro; IPR002314; tRNA-synt 2b.
CC InterPro; IPR006195; tRNA ligase II.
CC Pfam; PF03129; HGTP anticodon; 1.
CC Pfam; PF00587; tRNA-synt 2b; 1.
CC TIGRFAMs; TIGR00442; hisS; 1.
CC PROSITE; PS50862; AA tRNA LIGASE II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; Atp-binding;
CC Complete proteome.
SQ SEQUENCE 432 AA; 49183 MW; F6C8B0AD0BE150F CRC64;

```

```

Query Match 4.9%; Score 96; DB 1; Length 432;
Best Local Similarity 21.4%; Pred. No. 2.5;
Matches 61; Conservative 44; Mismatches 100; Indels 80; Gaps 14;

Qy 66 NMVAFDRMELDKVEPFEPFSPYHEALEPFPYMGQYIRPLVNFPSYGVNVSGVM 125
Db 77 DMTSSVRLVYNNM---FQTAP-----KPIKWY-----YIANMFYEEQSGYRPFQWA 122
Qy 126 EEOLKQDKVLLSNHQTEDADPAVIALMLETTPHISENIYVA-----GDRVIT 175
Db 123 GVELIGSDKY-----EADAEVIALFVES-----YLATGLREFTVNIIGDRILL 164
Qy 176 DPLCKPFSMGRN--LLCVYSKHMNDVPPELAEMKRSNTRSLKEMALLRGGSKIWIAP 233

```

```

Db 165 DEFAMLVGVDDIGIMRIIDKKO-----KLPQEDFINALKEFGDKNGIEKVELIN 216
Qy 234 SGGDRDPD---PITNWFPPAPFDATSLDNMRRLVD-----HAGLV-CHIVPL 276
Db 217 IKG--KPEVLPLAEELFTSEVAKNEINRLVALIDLLEAYEVKDWIIDJGIARGFDYIT 274
Qy 277 AILCHDIMPPLQVEKEIGKR---LISFHG-----TGISVAPE 312
Db 275 SIVFEATAPNDLGIGSIGGGGRYDNLIEVFGKPTPATGFAIGIE 319

```

RESULT 14

```

ID PODK RICPR STANDARD; PRT; 880 AA.
AC Q9ZDS5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyruvate,phosphate dikinase (EC 2.7.9.1) (Pyruvate,orthophosphate
DE dikinase).
GN PPK OR RP492.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -! FUNCTION: Catalyzes the reversible phosphorylation of pyruvate and
CC phosphate (By similarity).
CC -! CATALYTIC ACTIVITY: ATP + pyruvate + phosphate = AMP +
CC phosphoenolpyruvate + diphosphate.
CC -! COFACTOR: Magnesium ion (By similarity).
CC -! SUBUNIT: Homodimer (By similarity).
CC -! DOMAIN: The N-terminal domain contains the ATP/pi active site, the
CC central domain the pyrophosphate/phosphate carrier His-460, and
CC the C-terminal domain the pyruvate active site.
CC -! MISCELLANEOUS: The reaction takes place in three steps, each
CC mediated by a carrier histidine residue located on the surface of
CC the central domain. The two first partial reactions are catalyzed
CC at an active site located on the N-terminal domain, and the third
CC partial reaction is catalyzed at an active site located on the C-
CC terminal domain. For catalytic turnover, the central domain
CC swivels from the concave surface of the N-terminal domain to that
CC of the C-terminal domain.
CC -! SIMILARITY: Belongs to the PEP-utilizing enzyme family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AJ235272; CAA14944.1; -
CC PIR; F71652; F71652.
CC HSP; P22983; 2DIX.
CC InterPro; IPR008279; PEP mobile.
CC InterPro; IPR000121; PEP utilizers.
CC InterPro; IPR002192; PPK N term.
CC Pfam; PF00391; PEP-utilizers; 1.
CC Pfam; PF02896; PEP-utilizers; 1.
CC Pfam; PF01326; PPK N term; 1.
CC ProDom; PD000940; PEP utilizers; 1.
CC PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; 1.

```

CYTOPLASMIC (POTENTIAL):

FT	TRANSMEM	71	91	POTENTIAL.
FT	DOMAIN	92	115	LUMENAL (POTENTIAL).
FT	TRANSMEM	116	135	POTENTIAL.
FT	DOMAIN	136	278	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	279	298	POTENTIAL.
FT	DOMAIN	299	327	LUMENAL (POTENTIAL).
FT	TRANSMEM	328	345	POTENTIAL.
FT	DOMAIN	346	786	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	787	806	POTENTIAL.
FT	DOMAIN	807	816	LUMENAL (POTENTIAL).
FT	TRANSMEM	817	837	POTENTIAL.
FT	DOMAIN	838	857	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	858	880	POTENTIAL.
FT	DOMAIN	881	950	LUMENAL (POTENTIAL).
FT	TRANSMEM	951	970	POTENTIAL.
FT	DOMAIN	971	983	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	984	1002	POTENTIAL.
FT	DOMAIN	1003	1017	LUMENAL (POTENTIAL).
FT	TRANSMEM	1018	1038	POTENTIAL.
FT	DOMAIN	1039	1061	CYTOPLASMIC (POTENTIAL).
FT	MOD RES	383	383	PHOSPHORYLATION (BY SIMILARITY).
FT	METAL	731	731	MAGNESIUM (BY SIMILARITY).
FT	METAL	735	735	MAGNESIUM (BY SIMILARITY).
FT	METAL	336	336	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	337	337	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	339	339	CALCIUM 2 (VIA CARBONYL OXYGEN) (BY SIMILARITY).
FT	METAL	341	341	CALCIUM 2 (BY SIMILARITY).
FT	METAL	797	797	CALCIUM 1 (BY SIMILARITY).
FT	METAL	800	800	CALCIUM 2 (BY SIMILARITY).
FT	METAL	825	825	CALCIUM 1 (BY SIMILARITY).
FT	METAL	828	828	CALCIUM 2 (BY SIMILARITY).
FT	METAL	829	829	CALCIUM 1 (BY SIMILARITY).
FT	METAL	829	829	CALCIUM 2 (BY SIMILARITY).
FT	METAL	961	961	CALCIUM 1 (BY SIMILARITY).
FT	CONFLICT	314	314	P -> H (IN REF. 2).
FT	CONFLICT	430	430	T -> M (IN REF. 2).
FT	CONFLICT	439	439	M -> S (IN REF. 2).
FT	CONFLICT	659	659	T -> I (IN REF. 2).
FT	CONFLICT	859	859	T -> P (IN REF. 2).
FT	CONFLICT	1046	1046	G -> V (IN REF. 2).
SQ	SEQUENCE	1061 AA;	116180 MW; 55B6126ESD539822 CRC64;	

Query Match 4.8%; Score 94; DB 1; Length 1061;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 54; Conservative 45; Mismatches 106; Indels 36; Gaps 10;

QY	11	IDARSEQDILL---	SGIQRELEAGTLPKHTAQAMEELYQYKNVAVLSQAAPHAEIVLSNM	67
Db	522	VDSGKGLLVKGA	VENVLERSTHIQLDGSTRQLDQYSDILQS---	LHDMLSAL 577
QY	68	RVAFDRMFLDVKEPF	-----EPSVHEALIEPFTNYMFGQYIRPLVNFRESYVGNVSV	121
Db	578	R-CLGFAYSDVSD	FATYDGSDEHPAHQQLNPSYSSIESNLV-----	FVGFVGL 627
QY	122	FGVMEQLKQGD	KVLISNHOEADPAVIALMLET-TNPHISENIYVAGDRVITDPLCK	180
Db	628	RDPPEKREVRQA	-----IADCR-----AGIRVMVITGDKSTAEACREIGVFEADEDISS	678
QY	181	PFSMGRNLLCVYSKK	HMNDVPELAEMKRSNTRSLKEMALLIRGSKIIWIAPSGGRDR	239
Db	679	RSUTGKEFMDVKDQ	KNHLRQTGGL--LFSRAEPKHKQEIIVLLKEDGEVAMTGDGVNDA	736
QY	240	P 240		
Db	737	P 737		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 16, 2004, 02:08:21 ; Search time 102 Seconds
(without alignments)
2034.820 Million cell updates/sec

Title: US-09-926-805-7

Perfect score: 1949

Sequence: 1 GSHGHSRTFIDARSQDL.....KGLASTPSVLSQPLQFLD 374

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US09926805/runat_11032004_141821_4079/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09926805 @CNC 1.1.69 @runat_11032004_141821_4079 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/6C COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1380	70.8	1380	6	Patent No. 5210189-1
2	1380	70.8	1445	1	Sequence 1, Appli
3	1380	70.8	1445	6	Patent No. 5210189-4
4	1361	69.8	1104	3	Sequence 3, Appli
5	1360	69.8	1104	3	Sequence 5, Appli
6	1358	69.7	1104	3	Sequence 9, Appli
7	1316	67.5	1104	3	Sequence 1, Appli
8	1305	67.0	1104	3	Sequence 7, Appli
9	982	50.4	3397	6	Patent No. 5210189-3
10	536	27.5	1230025	4	US-09-198-452A-1
11	103.5	5.3	1011	4	US-09-134-001C-2712
12	100.5	5.2	10061	4	US-09-221-017B-672

13	99	5.1	1257	3	US-08-844-054-1	Sequence 1, Appli
14	99	5.1	1257	3	US-09-347-333-1	Sequence 1, Appli
15	98	5.0	4965	4	US-08-961-527-143	Sequence 143, App
16	97	5.0	1281	4	US-09-134-000C-996	Sequence 996, App
17	96.5	5.0	1728	4	US-09-328-352-3595	Sequence 3595, Ap
18	95	4.9	5583	4	US-09-312-283C-372	Sequence 372, App
19	95	4.9	10348	2	US-08-457-273B-41	Sequence 41, Appl
20	95	4.9	10348	3	US-08-556-419-13	Sequence 13, Appl
21	95	4.9	10348	3	US-09-041-886-14	Sequence 14, Appl
22	94	4.8	5672	4	US-09-023-655-1392	Sequence 1392, Ap
23	94	4.8	10366	1	US-08-246-982A-5	Sequence 5, Appli
24	94	4.8	10366	1	US-08-453-265-5	Sequence 5, Appli
25	93.5	4.8	1362	4	US-09-252-991A-14697	Sequence 14697, A
26	92.5	4.7	1888	4	US-09-270-957-27	Sequence 27, Appl
27	92.5	4.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
28	92.5	4.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
29	92	4.7	4758	3	US-09-191-647-1	Sequence 1, Appli
30	92	4.7	4758	3	US-09-540-245A-1	Sequence 1, Appli
31	92	4.7	4758	3	US-09-540-153-1	Sequence 1, Appli
32	88.5	4.5	1209	4	US-09-252-991A-1340	Sequence 1340, Ap
33	88.5	4.5	1260	4	US-09-252-991A-1263	Sequence 1263, Ap
34	88.5	4.5	1572	4	US-09-252-991A-1220	Sequence 1220, Ap
35	88	4.5	1827	4	US-09-134-001C-558	Sequence 558, App
36	88	4.5	2118	4	US-09-328-352-3310	Sequence 3310, Ap
37	88	4.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
38	88	4.5	441529	3	US-09-103-840A-1	Sequence 1, Appli
39	87.5	4.5	2733	4	US-08-997-685A-1	Sequence 1, Appli
40	87.5	4.5	3025	4	US-09-976-594-552	Sequence 552, App
41	87.5	4.5	3760	4	US-09-976-594-213	Sequence 213, App
42	87.5	4.5	6765	2	US-08-677-010-2	Sequence 2, Appli
43	87.5	4.5	6765	2	US-08-790-519-2	Sequence 2, Appli
44	87.5	4.5	13086	4	US-08-956-171E-16	Sequence 16, Appl
45	87.5	4.5	80161	3	US-09-036-987A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
5210189-1
; Patent No. 5210189
; APPLICANT: MURATA, NORIO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
; 3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/587,676
; FILING DATE: 25-SEP-1990
; SEQ ID NO:1:
; LENGTH: 1380
5210189-1

Alignment Scores:
Pred. No.: 4.03e-162 Length: 1380
Score: 1380.00 Matches: 252
Percent Similarity: 84.8% Conservative: 57
Best Local Similarity: 69.23% Mismatches: 55
Query Match: 70.81% Indels: 0
DB: 6 Gaps: 0

US-09-926-805-7 (1-374) x 5210189-1 (1-1380)

Qy	6	HisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleClnArg	25
Db	280	CATTCCTCCCTACTTCTTCCTGGATGCCGGAAGTGAACAGATCTTTTATCTGGTATCAAGAA	339
Qy	26	GluLeuGluAlaGlyThrLeuProlyshisIleAlaGlnAlaMetGluGluLeuTyGln	45
Db	340	GAAGCTGAAGCTGGAAGCTTCCAGCAATGTTCCAGCAGGAATGGAAGATTTGTTGG	399
Qy	46	AsnTyrIlysAenAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer	65
Db	400	AACTACAAAAATGCAGTCTTTAAGTAGTGAGCTTCCAGGCAGCATGAACCTGTTGTATCA	459

Qy	66	AsnVetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer	85
Db	460	AACATGTCTGTGTTTGTATCGCATGCTCTTGCTGTGGAGGATCCTTTATACITTTTAAT	519
Qy	86	ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle	105
Db	520	CCATATCATAAAGCAGTCAGAAACCAATTTCAGCTACTACATGTTGTGTCATACATACATC	579
Qy	106	ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet	125
Db	580	CGTCTCTTATGTATTCAAAAATTCGTACGTTGGAAATGCTTCTATATCTCTGAGCTG	639
Qy	126	GluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla	145
Db	640	GAACACAGATTTCACAGGGACACAATATCGTGTGTATATCAAAACCATCAAAGTGAAGCT	699
Qy	146	AspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIle	165
Db	700	GATCCGGCTGTCAITTCCTATTCCTTGAAGCACAAATCTCTCTTCATGAGAGAGAATC	759
Qy	166	IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly	185
Db	760	AAATGTGTGGCTGTGTATCGAGTCATCCTGATCTCTTTGTAAACGGTTCAGTATGGGA	819
Qy	186	ArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGlu	205
Db	820	AGGAACCTCATATGTTTACTCGAAAAAGCACATGAATGTTGATCTCGAGCTGTTGAC	879
Qy	206	MetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGly	225
Db	880	ATGAAAGAAAGCAACACACAGAGCTTAAAGAGATGGCTACAATGCTAAAGTCTGGC	939
Qy	226	SerLysIleIleTrpIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsn	245
Db	940	GGTCAACTTATATGGATTGCACCAAGCGGTGGAAAGGACCGCCCGAATCTCTTACTGGG	999
Qy	246	GlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp	265
Db	1000	GAATGGTTCCTGCACCCCTTTGATGCTTCTCGTAGACAACATGAGAAAGCTGGTTGAA	1059
Qy	266	HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro	285
Db	1060	CAITCTGGCGCTCCTGGACATATATATCCCAATGTCTTTGTGTTGCTATGATGACATGCC	1119
Qy	286	ProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPheHisGlyThr	305
Db	1120	CCTCCACCCAGGTTGAGAAAGAAATCGGAGAGAAACATTAGTTGGGTTTCACGGTACT	1179
Qy	306	GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro	325
Db	1180	GGACTATCAATTGCTCTGAAATCAACTTCTCAGACGTCACAGCAGACTCGAGAGCCCT	1239
Qy	326	GluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys	345
Db	1240	AATGAGCGGAAAGAGCATACGCCAAGCTTTGTACAAGTCGGTGAATGAACAAATACGAG	1299
Qy	346	ValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrProSerValSer	365
Db	1300	ATCTTAAACTCTCGGATTAACACAGNAGGAGTAGAAGCATCACTTCAAGGGTCTCT	1359
Qy	366	LeuSerGlnPro	369
Db	1360	TTGTCAACACT	1371

1 CORRESPONDENCE ADDRESS:
2 ADDRESSEE: Foley & Lardner
3 STREET: 3000 K Street, N.W., Suite 500
4 CITY: Washington, D.C.
5 COUNTRY: USA
6 ZIP: 20007-5109
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: PatentIn Release #1.0, Version #1.2.25
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/087,732
14 FILING DATE: 16-JUL-1993
15 CLASSIFICATION: 800
16 PRIOR APPLICATION DATA:
17 PRIOR APPLICATION NUMBER: PCT/JP92/00024
18 FILING DATE: 14-JAN-1992
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: JP 15883/1991
21 FILING DATE: 16-JAN-1991
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: JP 283807/1991
24 FILING DATE: 04-OCT-1991
25 ATTORNEY/AGENT INFORMATION:
26 NAME: BENT, Stephen A.
27 REGISTRATION NUMBER: 29,768
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (202)672-5300
30 TELEFAX: (202)672-5399
31 TELEX: 904136
32 INFORMATION FOR SEQ ID NO: 1:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 1445 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37 TOPOLOGY: linear
38 MOLECULE TYPE: cDNA
39 FEATURE:
40 NAME/KEY: CDS
41 LOCATION: 16..1392
42 FEATURE:
43 NAME/KEY: sig_peptide
44 LOCATION: 16..285
45 FEATURE:
46 NAME/KEY: mat_peptide
47 LOCATION: 286..1392
48 US-08-087-732-1

```
Db 475 AACATGTCTGTGCTTTTGGATCGCATGCTCTTGGTGTGGAGATCCTTTATACCTTTAAT 534
Qy 86 ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 535 CCATATCATTAAGCAGTCAGAACCAATTTGACTACTACATCTTTGTTCATACATACATC 594
Qy 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 595 CGTCTCTTATTGATTCACAAATTCGTACGTGTGAAATGCTTCTATATCTCTCAGCTG 654
Qy 126 GluGluGlnLeuLysGlyAspLysValValLeuLysSerAsnHisGlnThrGluAla 145
Db 655 GAAGCAAGATTCGACAGGACACCAATATCGGTGTTGATCAACCAATCAAAAGTGAAGCT 714
Qy 146 AspProAlaValIleAlaLeuMetLeuGluThrAsnProHisIleSerGluAsnIle 165
Db 715 GATCCGGCTGTCATTCCTATTGCTTGAAGCACAATCTCTTTCATAGGAGAACAT 774
Qy 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 775 AAATGTGTGGCTGTGATCGATCATCTATGCTCTTGTGAGCCGTTTCAGTATGGGA 834
Qy 186 ArgAsnLeuLysAlaValSerGlnAlaLeuTyrAspSerValCysGluGlnTyrIys 345
Db 1255 AATGAGCGCAAGAGCATACAGCCAGCTTTGTACAAAGTCGGTGAATGAACATACGAG 1314
Qy 346 ValLeuHisSerAlaValHisGlyGlyGlyLeuGluAlaSerThrProSerValSer 365
Db 1315 ATCTTAACCTCGGATTAAACACAGAGAGAGTAGAAGCATCAACTTCAAGGGTCTCT 1374
Qy 366 LeuSerGlnPro 369
Db 1375 TTGTACACACCT 1386
```

RESULT 3

```
5210189-4
; Patent No. 5210189
; APPLICANT: MURATA, NORIO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
; 3-PHOSPHATE ACYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/587,676
; FILING DATE: 25-SEP-1990
; SEQ ID NO: 4:
```

j LENGTH: 1445

5210189-4

Alignment Scores:

Score:	4,35e-162	Length:	1445
Pred. No.:	1380.00	Matches:	252
Percent Similarity:	84.89%	Conservative:	57
Best Local Similarity:	69.23%	Mismatches:	55
Query Match:	70.81%	Indels:	0
DB:	6	Gaps:	0

US-09-926-805-7 (1-374) x 5210189-4 (1-1445)

```
Qy 6 HisSerArgThrPheIleAspAlaArgSerGluGlnAspIleLeuSerGlyIleGlnArg 25
Db 295 CATTCCTGCTCTCTTCTTGGATGCGCGAAGTGAAACAAGATCTTTTATCTGGTATCAAGAAG 354
Qy 26 GluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGln 45
Db 355 GAAGCTGAAGCTGGAAGGTTGCCAGCAAAATGTCAGCAGCAATGGAAGAAATGTTATGG 414
Qy 46 AsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer 65
Db 415 AACTACAAAATGCAAGTTTAAAGTAGTGGAGCTTCCAGGGCAGATGAACTGTTGTATCA 474
Qy 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85
Db 475 AACATGTCTGTGCTTTGATTCGCATCTCTTGGTGTGGAGGATCCTTATACCTTTAAT 534
Qy 86 ProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIle 105
Db 535 CCATATCATTAAGCAGTCAGAACCAATTTGACTACTACATCTTTGTTCATACATACATC 594
Qy 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125
Db 595 CGTCTCTTATTGATTCACAAATTCGTACGTGTGAAATGCTTCTATATCTCTCAGCTG 654
Qy 126 GluGluGlnLeuLysGlyAspLysValValLeuLysSerAsnHisGlnThrGluAla 145
Db 655 GAAGCAAGATTCGACAGGACACCAATATCGGTGTTGATCAACCAATCAAAAGTGAAGCT 714
Qy 146 AspProAlaValIleAlaLeuMetLeuGluThrAsnProHisIleSerGluAsnIle 165
Db 715 GATCCGGCTGTCATTCCTATTGCTTGAAGCACAATCTCTTTCATAGGAGAACAT 774
Qy 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185
Db 775 AAATGTGTGGCTGTGATCGATCATCTATGCTCTTGTGAGCCGTTTCAGTATGGGA 834
Qy 186 ArgAsnLeuLysAlaValSerGlnAlaLeuTyrAspSerValCysGluGlnTyrIys 205
Db 835 AGCAACCTCATATGCTGTTTACTCGAAGAACCAATGATGATGCTGCTGAGCTTGTGAC 894
Qy 206 MetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuArgGlyGly 225
Db 895 ATGAAAAGAAAAGCAACACACCAAGCTTAAAGAGAGTGGCTACAAATGCTAAAGTCTGCG 954
Qy 226 SerLysIleIleThrIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245
Db 955 GGTCAACCTTATATGGATTGCAACCAAGCGGTGGAAGGACCCCGCAATCTCTCTCTGGG 1014
Qy 246 GlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265
Db 1015 GAATGGTTTCTGCAACCTTTGATGCTTCTCGGTAGACAAATGAGAACCTGGTTGAA 1074
Qy 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285
Db 1075 CATTCCTGGCTCTCTGGACATATATATCAATGCTCTTGTGCTTGTATGACATCATGCC 1134
Qy 286 ProProLeuGlnValGluLysGluIleGlyGluLysArgLeuLeuSerPheHisGlyThr 305
Db 1135 CCTCCACCCAGGTTGAGAAAGAAATCGGAGAGAAATAGTGGGTTTTCAGGTTACT 1194
```

QY 306 GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro 325
 Db 1195 GGACTATCAATTCCTCTGGAATCAACTTCTCAGAGCTCACACGACTGCGAGACCTT 1254
 QY 326 GluGluAlaLysAlaIleTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys 345
 Db 1255 AATGAGCGGAAAGACATACAGCCCAAGCTTTGTACAAGTCGGTGAATGAACAATACGAG 1314
 QY 346 ValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrProSerValSer 365
 Db 1315 ATCTTAACCTCUCGATTAAACACAGAGGAGTAGAAGCATCAACTTCAAGGGTCTCT 1374
 QY 366 LeuSerGlnPro 369
 Db 1375 TTGTCACAACT 1386

RESULT 4

US-09-000-092-3
 ; Sequence 3, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihito
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1104 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1104
 ; US-09-000-092-3

Alignment Scores:
 Pred. No.: 6,54e-160 Length: 1104
 Score: 1361.00 Matches: 250
 Percent Similarity: 85.56% Conservative: 64
 Best Local Similarity: 68.12% Mismatches: 53
 Query Match: 69.83% Indels: 0

DB: 3 Gaps: 0
 US-09-926-805-7 (1-374) x US-09-000-092-3 (1-1104)
 QY 3 MetHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGly 22
 Db 1 ATGGCTAGCCACTCCGCAAAATTTCTCGATGTTCGCTCTGAAGAAGAGTTGCTCTCTGC 60
 QY 23 IleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGlu 42
 Db 61 ATCAAGAAGGAACAAGAGCTGGAAGCTGCCCTCAAAATGTTGCTGCGAGAAATGAAGAA 120
 QY 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
 Db 121 TTGTATCAGAAATATAGAAATGCTGTTATGAGATGGAAATCCAAAGCGATGAAT 180
 QY 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
 Db 181 GTTCTGTCTAACATGACTGTTGCAATTAGATCGCATATTTGTTGGATGTGGAGGATCCTTT 240
 QY 83 GluPheSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGln 102
 Db 241 GTCTTCTCATCACACCAAGCAATTCGAGAGCCCTTTTGATTACTATACATTTTGGCCAG 300
 QY 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
 Db 301 AACTATATACGGCAATTCATGATTTTGGAAATTCATTCGTTGGTAACCTTCTCTTTTC 360
 QY 123 GlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGln 142
 Db 361 AAGGATATAGAAGAGAGAGCTTAAGCAGCGGTCAACACATCATCTTAATCCAAACCATCAA 420
 QY 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
 Db 421 AGTGAAGCAGATCCCGCAGTGTTCATTACTTCTGGAGAGACAAATTCATAATCCCA 480
 QY 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPhe 182
 Db 481 GAAACTTGTATCATACAGAGGTGATCGAGTTATTAACAGATCCTCTTTGACAGCCCTTT 540
 QY 183 SerMetGlyArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGlu 202
 Db 541 AGCATGGGAAGAAATCTTTTGTCTTACTTAAGAAGCACATGTATGATGATCCCGAG 600
 QY 203 LeuAlaGluMetLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeu 222
 Db 601 CTGTGTGATGAAGAAAGAAAGCAAAATACAGAGTTTGAAGAGTTGGTCTTACTTTA 660
 QY 223 ArgGlyGlySerLysIleIleTrpIleAlaProSerGlyGlyArgAspArgProAspPro 242
 Db 661 AGAGTGGTTCAAAAAATAATCTGGATTGCCACCCAGTGTGGAAGAGATCGTCCAGATGCT 720
 QY 243 IleThrAsnGlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
 Db 721 GTCACTGGTGAATGGTACCCAGCACCCCTTTGATGCTTCTTCAGTGGCAACATGAGAAG 780
 QY 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAsp 282
 Db 781 CTTATTCACAAATTCGATGTTCTCGGCAATTTGTTCCCTTCTTATTATGTCATGAC 840
 QY 283 IleMetProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPhe 302
 Db 841 ATCATGCCCTCCCTCACAGTTCGAAATTCGAGAAATTCGAGAAATTCGAGAAATTCGAG 900
 QY 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
 Db 901 AATGGCGCGGTTTGTCTGTGGCTCCCTGAAATTCAGCTTCGAGGAAATTCGCTACCCAC 960
 QY 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
 Db 961 AAAAATCTGAGAGGTTAGGAGGCATACACTCAAGGCATCTGTTGATTCTGTGCCCATG 1020
 QY 343 GlnTyrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrPro 362

Db 1021 CAATCAATGTGTCTCAAAACGGCTATCTCCGCAACAGGACTAGGAGCTTCAACTGCG 1080
 Qy 363 SerValSerLeuSerGlnPro 369
 Db 1081 GATGCTCTTTTGTCAACCT 1101

RESULT 5

US-09-000-092-5
 ; Sequence 5, Application US/09000092
 ; Patent No. 6160203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/000,092
 ; FILING DATE: 26-JAN-1998
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/JP96/01844
 ; FILING DATE: 03-JUL-1996

PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 192123/1995
 ; FILING DATE: 27-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/916

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1104 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1104

US-09-000-092-5

Alignment Scores:

Pred. No.: 8 72e-160 Length: 1104
 Score: 1360.00 Matches: 250
 Percent Similarity: 85.56% Conservative: 64
 Best Local Similarity: 68.12% Mismatches: 53
 Query Match: 69.78% Indels: 0
 DB: 3 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-5 (1-1104)

Qy 3 MethHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuSerGly 22
 Db 1 ATGGCTAGCCACTCCGCAAAATTTCTCGATGTCGCTCGAGAGAGTGTCTCTCTGCG 60
 Qy 23 IleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGlu 42

Db 61 ATCAAGAGGAAACAGAAAGCTGGAAGCTGCTCCAAATGTTGCTGCGAAGATGGAAGAA 120
 Qy 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
 Db 121 TTGTATCAGAAATTAGAAATGCTGTATTAGAGTGGAAATCCAAAGCGCATGAAT 180
 Qy 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
 Db 181 GTTCTGTCTAAACATGACTGTTGCATTAGATCGCATATTGTTGGATGTGGAGTGCCTTTT 240
 Qy 83 GluPheSerProTyrHisGluAlaIleLeuGlnProPheAsnTyrTyrMetPheGlyGln 102
 Db 241 GTCCTTCTCATCACCAACAGCAATTCGAGAGCTTTTGATTACTATGATCTTTTGGCCAG 300
 Qy 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
 Db 301 AACTATATACGGCCATTGATTGATTTTGGAAATTCATTGCTTGGTAACTTCTTCTTTC 360
 Qy 123 GlyValMetGluGlnGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGln 142
 Db 361 AAGGATATAGAGAGAGAGCTTAAAGCAGGCTGACACATCATCTTAATGTCCAACCATCAA 420
 Qy 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
 Db 421 AGTGAAGCAGATCCGCGAGTGTATGCTTCTTGGAGAGACAAATCTCTAATCCGA 480
 Qy 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuLysLysProPhe 182
 Db 481 GAAAACCTTGATCTACATAGCAGGTGATCGAGTTATTAACAGATCCTCTTTGCAAGCCCTT 540
 Qy 183 SerMetGlyArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGlu 202
 Db 541 AGCATGGAGAGGAATCTCTTGTGTTTACTCTAAGAAGCACATGTATGATGATCCCGAG 600
 Qy 203 LeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeu 222
 Db 601 CTCACAGAAACAAAAGGAAGCAACACAGAGTCTTAAGGAGAGTGGCTTACTCTTA 660
 Qy 223 ArgGlySerLysIleIleTyrPheAlaProSerGlyArgAspArgProAspPro 242
 Db 661 AGAGTGGATCACAACTAATATGATGATTCACCCAGTGTGGTGGGACCGCGCGATCCC 720
 Qy 243 IleThrAsnGlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
 Db 721 TCGACTGGAGATGTTACCCAGCACCTTTGATGCTTCTTCAGTGGACAAACATGAGAGG 780
 Qy 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAsp 282
 Db 781 CTTATTCAACATTCGGATGTTCTCGGCAATTTGTTTCCCTTGTCTTATTATGTCATGAC 840
 Qy 283 IleMetProProProLeuGlnValGluLysGluIleGlyLysArgLeuIleSerPhe 302
 Db 841 ATCATGCCCTCCCTCCACAGGTGCAAAATTTGAAATTCAGTTCGAGGAAATGTGTACCAC 900
 Qy 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
 Db 901 AATGGGCGGGTGTCTGTGGCTCTGAAATTCAGTTCGAGGAAATGTGTACCAC 960
 Qy 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
 Db 961 AAAAATCTCGAGAGGTTAGGAGGCATCTCAAGGACATGTTTGTATTCTGTGGCCATG 1020
 Qy 343 GlnTyrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrPro 362
 Db 1021 CAATACAATGTCTCAAAACGGCTATCTCCGCAAAACAGGACTAGGAGTTCAACTGCG 1080
 Qy 363 SerValSerLeuSerGlnPro 369
 Db 1081 GATGCTCTTTTGTCAACCT 1101

RESULT 6

US-09-000-092-9

Sequence 9, Application US/09000092
 Patent No. 6160203
 GENERAL INFORMATION:
 APPLICANT: FERRI, Stefano
 APPLICANT: TOGURI, Toshihiro
 TITLE OF INVENTION: DNA STRANDS CODING FOR
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/000,092
 FILING DATE: 26-JAN-1998
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP96/01844
 FILING DATE: 03-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 192123/1995
 FILING DATE: 27-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16887/916
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1104
 US-09-000-092-9

Alignment Scores:
 Pred. No.: 1104
 Score: 1358.00
 Percent Similarity: 85.99%
 Best Local Similarity: 68.41%
 Query Match: 69.68%
 DB: 3
 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-9 (1-1104)

QY 6 HisSerArgThrPheLeuAlaValLeuGlnAspLeuLeuSerGlyIleGlnArg 25

DB 10 CACTCCCGCAAAATTCGATGTTGCTGAGAGAGAGTTGCTCTCTCGCATCAAGAAG 69

QY 26 GluLeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGln 45

DB 70 GAAACAGAGCTGGAAGTCTCCAAATGTTGCTGAGGAGTGAAGATTGATCAG 129

QY 46 AsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSer 65

DB 130 AATTATAGAAATGCTGTTATTGAGAGTGAATCCAAAGGCAGATGAAATTTCTGTCT 189

QY 66 AsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSer 85

DB 1104

DB 190 AACATGACTGTTGCATTAGATCGCATATATGTTGGATGTGGAGGATCCCTTTTGTCTCTCA 249

QY 86 ProTyrHisGluAlaIleLeuGluProPheLeuTyrMetPheGlyGlnAsnTyrIle 105

DB 250 TCACACCACAAAGCAATTCGAGAGCCTTTGATTACTACATTTTGGCCAGCAATATATA 309

QY 106 ArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMet 125

DB 310 CGGCATTGATTGTTGGAAATTCATGTTGTAACCTTTCTCTTTCAAGATATA 369

QY 126 GluGluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAla 145

DB 370 GAAGAGAACTTAAGCAGGTTGACCAACATCATCTTATGTCCACCATCAAGGAACA 429

QY 146 AspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIle 165

DB 430 GATCCCGCAGTGATTGCAATTTCTTGGAAGAAGCAAAATTCACATAATCGCAAAACATTG 489

QY 166 IleTyrValAlaGlyAspArgValIleThrAspProLeuCysLysProPheSerMetGly 185

DB 490 ATCTACATAGCAGGTGATCGAGTTATAACAGATCTCTTTGCAAGCCCTTTAGCATGGA 549

QY 186 ArgAsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGlu 205

DB 550 AGGAATCTTCTTTGTTTACTCTTAAGAAGCACATGTATGATGATCCGAGCTTGTTCAT 609

QY 206 MetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGly 225

DB 610 GTAAGAAAGAGCAAAATACAGAGTTTGAAGAGTTGGTCTTACTTTTAAGAGGTGGT 669

QY 226 SerLysIleIleTyrPheAlaProSerGlyGlyArgAspArgProAspProIleThrAsn 245

DB 670 TCAAAAATAATCTGATTTGACCCAGTGGTGAAGAGATCTCCAGATGCTGCTACTGGT 729

QY 246 GlnTrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAsp 265

DB 730 GAATGGTACCAGCACCCCTTTGATGCTTCTTCAGTGGACACATGAGAGGCTTATTCAA 789

QY 266 HisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 285

DB 790 CATTGGATGTTCTGGCAATTTGTTCCCTTCTTATTTATGTCATGACATATGCC 849

QY 286 ProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPheHisGlyThr 305

DB 850 CTTCCCTCACAGGTCGAAATTTGAATTTGGAGAAAAAGAGTGTGCTTTAATGGGCG 909

QY 306 GlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerPro 325

DB 910 GGTGTGTGTGGCTCTCTGAAATCAGCTTCGAGGAAATTTGCTGTCTACCCCAAAAATCT 969

QY 326 GluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLys 345

DB 970 GAGGAGTTAGGAGGATCTCAAGGACACTGTTGATTTCTGGCCATCAATACAT 1029

QY 346 ValLeuHisSerAlaValHisGlyGlyGlyGlyLeuGluAlaSerThrProSerValSer 365

DB 1030 GTGCTCAAAACGGTATCTCCGCAACCAAGACTTAGAGCTTCAACTCGGAGTCTCT 1089

QY 366 LeuSerGlnPro 369

DB 1090 TTGTCACAACT 1101

RESULT 7

US-09-000-092-1

; Sequence 1, Application US/09000092

; Patent No. 6160203

; GENERAL INFORMATION:

; APPLICANT: FERRI, Stefano

; APPLICANT: TOGURI, Toshihiro

; TITLE OF INVENTION: DNA STRANDS CODING FOR

; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/000,092
 FILING DATE: 26-JAN-1998
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP96/01844
 FILING DATE: 03-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 192123/1995
 FILING DATE: 27-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16887/916
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1104
 US-09-000-092-1

Alignment Scores:
 Pred. No.: 2,72e-154 Length: 1104
 Score: 1316.00 Matches: 243
 Percent Similarity: 83.11% Conservative: 62
 Best Local Similarity: 66.21% Mismatches: 62
 Query Match: 67.52% Indels: 0
 DB: 3 Gaps: 0

US-09-926-805-7 (1-374) x US-09-000-092-1 (1-1104)

```

QY 3 MethHisGlyHisSerArgThrPheIleAspAlaArgSerGluGlnAspLeuSerGly 22
Db 1 ATGGCTAGCCACTCTCGACATTCGTAACGTTCTGTCGAGAGAGCTGATATCTGAA 60

QY 23 IleGlnArgGluLeuGluAlaGlyThrLeuProTyrHisIleAlaGlnAlaMetGluGlu 42
Db 61 ATAAAAGGGAATCAGAAATGGAAGGTTACCTAAAGCTGTGCTATGCTATGAGGGA 120

QY 43 LeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIle 62
Db 121 CTTTTCCTACTACTCGCAATGTCAGTCTCTTTCAAGTGAATTTCTCATGCTGATGAATA 180

QY 63 ValLeuSerAsnMetArgValAlaPheAspArgMetPheLeuAspValLysGluProPhe 82
Db 181 GTGTGTCAACATGAGTGTATGCTGATTTGTTTGTGGATATTGAGGACCCCTTTT 240

QY 83 GluPheSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGln 102
Db 241 GTATTTCACCGGTTTCAACAGCTATTTCGAGAGCTCTGCTGACTATTATCTCTTGGTCAA 300

QY 103 AsnTyrIleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPhe 122
Db 301 GATTACATTGCGCCATTGGTAGATTTTGGAAATTCATATGTTGTAACATCGCCATTTC 360

```

```

QY 123 GlyValMetGluGluGlnLeuLysGlnGlyAspYsValValLeuIleSerAsnHisGln 142
Db 361 CAAGAAATGAGGAGAGGCTTAAGCAGGGTGACACATCATCTTAATGTCACCAACCATCA 420

QY 143 ThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSer 162
Db 421 AGTGAAGCAGATCCCGCAGTGAATGCTTCTTGGAGAGACAAATTCATAATCGCA 480

QY 163 GluAsnIleIleTyrValAlaGlyAspArgValIleThrAspProLeuLysCysIleSer 182
Db 481 GAAACTTGATCTACATACAGCAGTGAATGCTTCTTGGAGAGACAAATTCATAATCGCA 480

QY 183 SerMetGlyArgAsnLeuLysCysValTyrSerLysLysHisMetAsnAspValProGlu 202
Db 541 AGCATGGGAGGAATCTCTTTGTGTTTACTCTAAGAAGCACATGATGATGATGATGATG 600

QY 203 LeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeu 222
Db 601 CTTGTTGATGTAAAGAAAGAGCAATAACAAGGAGTTTGAAGAGTGTGCTTCTTACTTTA 660

QY 223 ArgGlyGlySerLysIleIleTyrIleAlaProSerGlyGlyArgAspArgProAspPro 242
Db 661 AGAGTGGTTCAAAAATAATCTGGATTGCACCCAGTGGTGAAGAGATCGTCCAGATGCT 720

QY 243 IleThrAsnGlnTyrPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArg 262
Db 721 GTCACTGGTGAATGGTATACCCAGCACCTTTGATGCTTCTTCAAGTGGACACATGAGAAG 780

QY 263 LeuValAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAsp 282
Db 781 CTTATTCAACATTCGGATGTTCTGGGCAATTTGTTCCCTTGTCTTATTAATGTCATGAC 840

QY 283 IleMetProProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPhe 302
Db 841 ATCATGCCCTCCCTCACAGTTCGAATTTGAATTTGGAGAAAGAGTGTGCTTGCCTTT 900

QY 303 HisGlyThrGlyIleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCys 322
Db 901 AATGGGGGGGTTTGTCTGTGGCTCTCTGAAATTCAGTTCGAGGAAATTTGCTTACCCAC 960

QY 323 GlySerProGluGluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGlu 342
Db 961 AAAAATCTGAGGAGGTAGGAGGCATCTCAAGGACCTGTTTGTGATCTGTGGCCATG 1020

QY 343 GlnTyrLysValLeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrPro 362
Db 1021 CAATACAATGTGCTCAAAACGGCTATCTCCGGCAACACAGGACTAGGAGCTTCAACTCG 1080

QY 363 SerValSerLeuSerGlnPro 369
Db 1081 GATGCTCTCTTTGTCAACACCT 1101

```

RESULT 8

US-09-000-092-7
 ; Sequence 7, Application US/09000092
 ; Patent No. 616203
 ; GENERAL INFORMATION:
 ; APPLICANT: FERRI, Stefano
 ; APPLICANT: TOGURI, Toshihiro
 ; TITLE OF INVENTION: DNA STRANDS CODING FOR
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: FOLEY & LARDNER
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/000,092
 FILING DATE: 26-JAN-1998
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP96/01844
 FILING DATE: 03-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 192123/1995
 FILING DATE: 27-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16887/916
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1104 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1104
 US-09-000-092-7

Alignment Scores:
 Pred. No.: 6,43e-153 Length: 1104
 Score: 1305.00 Matches: 241
 Percent Similarity: 83.47% Conservative: 62
 Best Local Similarity: 66.39% Mismatches: 60
 Query Match: 66.96% Indels: 0
 DB: 3 Gaps: 0
 US-09-926-805-7 (1-374) x US-09-000-092-7 (1-1104)

QY 7 SerArgThrPheIleAspAlaArgSerGluGlnAspLeuLeuSerGlyIleGlnArgGlu 26
 Db 13 TCTCGCACTATCGTAACGTTCCTCGAGAGAGCTGATATCTCAAAATAAAGGGAA 72
 QY 27 LeuGluAlaGlyThrLeuProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsn 46
 Db 73 TCAGAAATGGAGGTACTTAAAGTGTGCTTATGCTATGGAGGACTTTTCACTAC 132
 QY 47 TyrLysAsnAlaValLeuGlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsn 66
 Db 133 TATCGCAATGCGAGTCTCTTCAAGTGAATTTCTCATGCTGATGAATAGTGTGTCAAC 192
 QY 67 MetArgValAlaPheAspArgMetPheLeuAspValLysGluProPheGluPheSerPro 86
 Db 193 ATGAGTGTATGCTGATTTGTTTGTGGATATGAGGACCTTTGTATTTCCACCG 252
 QY 87 TyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIleArg 106
 Db 253 TTTCAAAAGCTATTCGAGAGCGCTGCTACTATTATTCCTTTGGTCAAGATTACATTCCG 312
 QY 107 ProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyValMetGlu 126
 Db 313 CCAATGGTGAATTTGGAAATTCATATGTTGGTAACATGCCATTTTCCAAAGAAATGGAG 372
 QY 127 GluGlnLeuLysGlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAlaAsp 146
 Db 373 GAGAAGCTTAACGAGGGTGACACATCATCTTAATGTCACCATCAAGTGAAGACAT 432
 QY 147 ProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIleIle 166
 Db 433 CCGCGAGTGTATGCTACTTCTGAGAGAGACAAATTCATTAATTCGAGAAACTTGTATC 492

QY 167 TyrValAlaGlyAspArgValIleThrAspProLeuCysIlyBProPheSerMetGlyArg 186
 Db 493 TACATAGCAGGTGATCGAGTTATAACAGATCTCTTTGCAAGCCCTTTAGCATGGGAAG 552
 QY 187 AsnLeuLeuCysValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGluMet 206
 Db 553 AATCTTCTTTGGTTTACTCTAAGAAGCACATGTATGATGATCCCGAGCTTGTGTATGTA 612
 QY 207 LysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGlySer 226
 Db 613 AAGAAAGAGCAATAACAAGAGATTGTAAGAGTTTGGTCTTACTTTTAAAGAGGTGTTCA 672
 QY 227 LysIleIleTyrIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsnGln 246
 Db 673 AAAATATCTGGATTCACCCAGTGTGGAAGAGATCGTCCAGATGCTGTCACTGTGTAA 732
 QY 247 TrpPheProAlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAspHis 266
 Db 733 TGGTACCCAGCACCTTTGATGCTTCTTCAGTGGACAAACATGAGAAGGCTTATTCACAT 792
 QY 267 AlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetPro 286
 Db 793 TCGAGTGTTCCTGGGCAATTTTCCCTTGTCTTTATTTATGTATGTCACATCATGCCCT 852
 QY 287 ProLeuGlnValGluLysGluIleGlyGluLysArgLeuIleSerPheHisGlyThrGly 306
 Db 853 CCTCACAGTTCGAATTTGAATTTGAGAAAAGAGTGTGCTTTTATGGGGCGGT 912
 QY 307 IleSerValAlaProGluIleAsnPheGlnGluValThrAlaSerCysGlySerProGlu 326
 Db 913 TTGTCTGTGGTCTCTCAAAATTCAGCTTCGAGGAATTTGCTGTACCACAAAAATCTGTAG 972
 QY 327 GluAlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLysVal 346
 Db 973 GAGGTAGGAGGCATACCTCAAGGCACTGTGTGATTTGTGGCATGCAATACAAATGTG 1032
 QY 347 LeuHisSerAlaValHisGlyGlyLysGlyLeuGluAlaSerThrProSerValSerLeu 366
 Db 1033 CTCAAACGGCTATCTCCGGCAACAGGACTTAGGAGCTTCACTGCGGATGCTCTTTG 1092
 QY 367 SerGlnPro 369
 Db 1093 TCACAACCT 1101
 RESULT 9
 5210189-3
 ; Patent No. 5210189
 ; APPLICANT: MURATA, NORIO
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING GLYCEROL
 ; 3-PHOSPHATE ACYLTRANSFERASE
 ; NUMBER OF SEQUENCES: 4
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/587,676
 ; FILING DATE: 25-SEP-1990
 ; SEQ ID NO: 3
 ; LENGTH: 3397
 5210189-3
 Alignment Scores:
 Pred. No.: 8,95e-112 Length: 3397
 Score: 982.00 Matches: 243
 Percent Similarity: 43.54% Conservative: 57
 Best Local Similarity: 35.27% Mismatches: 52
 Query Match: 50.38% Indels: 340
 DB: 6 Gaps: 11
 US-09-926-805-7 (1-374) x 5210189-3 (1-3397)
 QY 18 AspLeuLeuSerGlyIleGlnArgGluLeuGluAlaGlyThrLeuProLysHisIleAla 37
 Db 1122 GATCTTTATCTGTATCTCAAGAGAGCTGAAGCTGAGAGTTGCCAGCAATGTTGCA 1181
 QY 38 GlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeuGlnSer----- 54

1182 GCAGGAATGGAAGAAATGTTATTGGAAGTACAAAAATGCAAGTAAATGAAATCCTTTTATAC 1241
QY |||||
54 ----- 54
1242 TCACATAGTTTACAGATTAGTCTAGCTTANGCTTTGGAAATTAATAGATTAAAGAACATCAGT 1301
Db |||||
QY -----AlaAla 56
1302 CCACTGAACCTTTGGCATTGTGCTTACTGATGATGGTTTATAGCTTTTAAAGTAGTGAGCTT 1361
Db |||||
QY ProHisAlaGluAspIleValLeuSerAsnMetArgValAlaPheAspArgMet-PheLeu 76
1362 CCAGG-GCAGATGAAACTGTGTATCAACATGCTGCTGCTTTGATCGCATGCTCTT 1420
Db |||||
QY u-----AspVal-- 78
1421 GGTGTGGAGGTACTTCTTCTGCAATCTCTTGGAAATGATTTTCTACTAGGAGATGTCAT 1480
Db |||||
QY LysGluProPheGluPh 84
1481 CTAAAAATCTTTTACTTGAGCTTCTCGCTGTGAATATTTTCAGGATCCTTATATCTT 1540
Db |||||
QY eSerProTyrHisGluAlaIleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTy 104
1541 TAATCCATATCATAAAGCAGTACAGAGAACATTTGACTACTACATGTTTGTCCATACATA 1600
Db |||||
QY rIleArgProLeuValAsnPhe----- 111
1601 CATCCGCTCTCTTATTGA-TTTCAAGTAAGCTGGAATATGCACTCATAGTTATCAGATTT 1659
Db |||||
QY ----- 111
1660 TATTCATAACTGAATGTAATTAGAAAGCAATATCGTTTGGTTAGTATCATGGATCTTCCA 1719
Db |||||
QY -----ArgG1 113
1720 GTTTTCTTCTGCAATCAAGTGTTCATCTACTAAACAACCTTTGAATATGTTATGCAGAA 1779
Db |||||
QY uSerTyrValGlyAsnValSerValPheGlyValMetGluGluGlnLeuLysGln- 131
1780 TTCGTACGTTGGAATGCTTCTATATCTCTGAGCTGGAAGAACAGATTTCGACAGTCA 1839
Db |||||
QY ----- 131
1840 CTTCCCTCTTCTGTGCTCTAGTTAATGCTGTGTTCTTTCATCAACAGTTAAAGTTCTT 1899
Db |||||
QY GlyAspLys-ValValLeuIles 139
1900 TCATTCAGGATTTTAAATCACTCAACTGTTTAAATACAGGACACACATATCGTTGTGATAT 1959
Db |||||
QY exAsnHisGlnThrGluAlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnTy 159
1960 CAACCATCAAGTGAAGTGTATCCGGCTGTCTATTCTCTATTGCTTGAAGCACAAATCTC 2019
Db |||||
QY rHisIleSerGluAsnIleIle----- 166
2020 CTTTCATAGGAGAGAACATTGTGAGCCTTCGAGCCTTGTCTGAGCTATTATAGGCTATT 2079
Db |||||
QY TyrValAlaGlyAspArgValIleT 175
2080 GTCTCTGTTCTTATTGTTTCTTTGTTGTCAGAAATGTGTGGCTGTGATCGATCATCA 2139
Db |||||
QY hrAspProLeuCysLysProPheSerMetGly----- 185
2140 CTGATCTCTTGTAAAGCGCTTCAGTATGGGAAGGTATCAGGAGCTTTCATTGTCATGG 2199
Db |||||
QY ----- 185
2200 TATGCTACCTGATGGGCTGAAGTGAATCTTAGCACCTTTGTATGACTTCGGGATTTCTA 2259
Db |||||
QY -----ArgAsnLeuLeuCysValTyrSerIlyshHisMetAsnAspValProGluL 203
186 |||||

2260 CTGCTTTTCAGGAACCTCATATGTTGTTTACTCGAAAAAGCACATGAATGTTGATCCTGAGC 2319
QY |||||
2303 euAlaGluMetIlyshysArgSerAsnThrArgSerIleuIlyshGluMetAlaLeuLeuLeu- 222
Db |||||
2320 TTGTTGATGAAAGAAAGAAAGCAACACACGAAAGCTTAAAGGAGATGGCTCAATGCTAA 2379
QY ----- 222
2380 GGTTAATGGAATAAATAGCGAGTCTTCTTATTTGATCTCATAGGAACAGAAATTAAGG 2439
Db |||||
QY -----ArgGlyGlySerIlyshIleI 229
2440 AAAACATTGATGTTTAGTCTGCTAATGTTGATGAACCCCTGTAGGTCTGCGCGTCAACTTA 2499
Db |||||
QY leTIpIleAlaProSerGlyGlyArgAspArgProAspProIleThrAsnGlnTyrPheP 249
2500 TATGATTTGCACCAAGCGGTGAAGGAGCGCCCGAATCCTTCTACTGGGGAATGGTTTC 2559
Db |||||
QY ro----- 249
2560 CTGTAAGTTATTGATGGAATACAGAGATCTTATCTGAGTCTGTGTAGATATCAAGTACT 2619
Db |||||
QY -----AlaProPheAspAlaThrSerLeuAspAsnMetArgArgLeuV 264
2620 AATCTAGTAATGTTTCAGGCACCCCTTTGATGCTTCTCGTAGACACACATGAGAGACTGG 2679
Db |||||
QY aAspHisAlaGlyLeuValGlyHisIleTyrProLeuAlaIleLeuCysHisAspIleM 284
2680 TTGAACATCTGCGCCTCTCGACATATATATCCAAATGTCCTTGTCTTGTGTAAGATCA 2739
Db |||||
QY etProProProLeu----- 288
2740 TGCCCCCTCCACC-CCAGGTATTGATGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 2798
Db |||||
QY -----GlnValGluL 292
2799 GGTAACTCTGTTGACTTTCCTCTCATCATCATCTTCTGTCATCTGCGAGTTGAGA 2858
QY yGluIleGlyGlyLysArgLeuIleSerPheHisGlyThrGlyIleSerValAlaProG 312
2859 AAGAAATCGGAGAGAAAGATTAGTTGGTTTCACGGTACTGACTATCAATGCTCTCTG 2918
Db |||||
QY luIleAsnPheGlnGluValThrAlaSerCysGlySerProGluGlu----- 327
2919 AAATCACTTCTCAGACGTCACAGCAGACTGCGAGAGCCCTAATGAGGTCTGTCTGTAA 2978
Db |||||
QY ----- 327
2979 ACTCAGCACCAATGACTTTAGTTTCAAGATCGAGTAAATTCATCATCATATAATAACA 3038
QY -----AlaLysAlaAlaTyrSerGlnAlaLeuTyrAspSerValCysG 342
3039 TTCCTGTGATGCGAGCGGAAAGAGATATCAGCAAGCTTTGTACAAGTCGGTGAATG 3098
Db |||||
QY luGlnTyrLysValLeuHisSerAlaValHisGlyGlySerGlyLeuGluAlaSerThrP 362
3099 AACAATACAGATCTTAAACTCTGCGATTAAACACAGAGAGAGAGTAGAAGCATCAACTT 3158
Db |||||
QY roSerValSerLeuSerGlnPro 369
3159 CAAGGTCTCTTTGTCACAACT 3181
Db |||||
RESULT 10
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6859294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A

```
/ CURRENT FILING DATE: 1998-11-24
/ NUMBER OF SEQ ID NOS: 6849
/ SEQ ID NO 1
/ LENGTH: 1230025
/ TYPE: DNA
/ ORGANISM: Chlamydia pneumoniae
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(15000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (15001)..(30000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (30001)..(45000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (45001)..(60000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (60001)..(75000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (75001)..(90000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (90001)..(105000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (105001)..(120000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (120001)..(135000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (135001)..(150000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (150001)..(165000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (165001)..(180000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (180001)..(195000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (195001)..(210000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (210001)..(225000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (225001)..(240000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (240001)..(255000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (255001)..(270000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (270001)..(285000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (285001)..(300000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (300001)..(315000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (315001)..(330000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (330001)..(345000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (345001)..(360000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (360001)..(375000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (375001)..(390000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (390001)..(405000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (405001)..(420000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (420001)..(435000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (435001)..(450000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (450001)..(465000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (465001)..(480000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (480001)..(495000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (495001)..(510000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (510001)..(525000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (525001)..(540000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (540001)..(555000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (555001)..(570000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (570001)..(585000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (585001)..(600000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (600001)..(615000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (615001)..(630000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (630001)..(645000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (645001)..(660000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (660001)..(675000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (675001)..(690000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
```

```

/ LOCATION: (690001)..(705000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (705001)..(720000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (720001)..(735000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (735001)..(750000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (750001)..(765000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (765001)..(780000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (780001)..(795000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (795001)..(810000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (810001)..(825000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (825001)..(840000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (840001)..(855000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (855001)..(870000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (870001)..(885000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (885001)..(900000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature
/ LOCATION: (900001)..(915000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc feature

```

Alignment Scores:

```

Pred. No.:      8.09e-52      Length:      1230025
Score:          536.00      Matches:      116
Percent Similarity: 56.06%      Conservative: 69
Best Local Similarity: 35.15%      Mismatches: 115
Query Match:      27.50%      Indels:      30
DB:              4          Gaps:      6

```

US-09-926-805-7 (1-374) x US-09-198-452A-1 (1-1230025)

```

QY 32 LeuProlysisHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaVal 51
Db 1104887 CTCACAGAGCTTATATCAGAAATTTCTGTTTTTCAACAGATTACATCGATGAGCT 1104828

QY 52 Leu---GlnSerAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAla 70
Db 1104827 ACAAAGAGCTGAGCTGATGAGTGTCTATGTTTACATGGGTCAAGGTTATC 1104768

QY 71 PheAspArgMetPheLeuaspValLysGluProPheGluPheSerProTyrHisGluAla 90
Db 1104767 ATCAG-----GACTTGAAAGAACCCCTTATTTCCTCGTATCATATAAAAA 1104720

QY 91 IleLeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyrIleArgProLeuValAsn 110
Db 1104719 ATTCAGCTCTATAGATCTATTCGTTGAGTATAGATTTTTTCTCTAGTCATGAC 1104660

QY 111 PheArgGluSerTyrValGlyAsnValSerValPheGlyValMetGluGluGlnLeuLys 130

```

```

Db 1104659 GATAAAACCTCTCGAATTTTAAACCTACCCGCTTAAGAAATAGAGAAATCAITGCT 1104600
QY 131 GlnGlyAspLysValValLeuIleSerAsnHisGlnThrGluAlaAspProAlaValIle 150
Db 1104599 AGGGAGAGATAATGTGTTCTCTCTGCAATCACCACAGAGATGCGATCCACAACATCG 1104540
QY 151 AlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsnIleIleTyrValAlaGly 170
Db 1104539 TATTATGCTTAGGCAAGACGCTATCTGAGTTAATGAGAACATGATTTTGTAGTGA 1104480
QY 171 AspArgValIleThrAspProLeuCysLysProPheSerMetGlyArgAsnLeuLys 190
Db 1104479 GACCGAGTCATCTCTGATCCCTAGCAGCACTTATGAGTGGATGTGATTTATATGT 1104420
QY 191 ValTyrSerLysLysHisMetAsnAspValProGluLeuAlaGluMetLysLysArgSer 210
Db 1104419 ATTATTTCTAAGCGTCATATTGCCACTCCACAGAACTCCGGAGAGAAAGCTTCTTCAT 1104360
QY 211 AsnThrArgSerLeuLysGluMetAlaLeuLeuLeuArgGlyGlySerLysIleIleTyr 230
Db 1104359 ATCAGAAAGCATGCAATATTTAAAGACCTTGTAAATGAAGGAGGCAATTTATTATC 1104300
QY 231 IleAlaProSerGlyGlyArgAspArgProAspProIleThrAsnGlnTyrPheProAla 250
Db 1104299 GTAGCCCTTCAGAGAGTCTGCACAGAAAAAT---GCGAGAGGAAGATATATCCTTCA 1104243
QY 251 ProPheAspAlaThrSerLeuAspAsnMetArgArgLeuValAspHisAlaGlyLeuVal 270
Db 1104242 GAATTCCTCCAGAAAGCATGGAAGTGTTCGTTTATTATGCTAAGGCTTCTAATCAACG 1104183
QY 271 GlyHisIleTyrProLeuAlaIleLeuCysHisAspIleMetProProLeuGlnVal 290
Db 1104182 ACACATTTTATCCCTTCCTTTAAACCTATGACATTTCTCCCTCCCTCCGAAAAA 1104123
QY 291 GluLysGluIleGlyGlyLysArgLeuIleSerPheHisGlyThrGlyIleSerValAla 310
Db 1104122 GAGATGCAATTTGGGAGCAGCAGCGAGCAATTTCTTT-----GCT 1104084
QY 311 ProGlu---IleAsnPheGlnGluValThrAlaSerCysGlySerProGluGluAlaLys 329
Db 1104083 CTGTTTTCTTCACTTCGGA-----1104063
QY 330 AlaAlaTyrSerGlnAlaLeuTyrAspSerValCysGluGlnTyrLysValLeuHisSer 349
Db 1104062 -----GCAGAACTATTTTGTGATGATTTGTTCAAAAGAGGAACTTATTCATGCG 1104012
QY 350 AlaValHisGlyGlyLysGlyLeuGluAla 359
Db 1104011 GATAAACACGCCCAAGAACATTAAGAGCA 1103982

RESULT 11
US-09-134-001C-2712
; Sequence 2712, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2712
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2712

```

```
Alignment Scores:
Pred. No.: 0.0058 Length: 1011
Score: 103.50 Matches: 49
Percent Similarity: 43.08% Conservative: 35
Best Local Similarity: 25.13% Mismatches: 78
Query Match: 5.31% Indels: 33
DB: 4 Gaps: 9

US-09-926-805-7 (1-374) x US-09-134-001C-2712 (1-1011)
Qy 49 AsnAlaValLeuGlnSerAla---AlaProHisAlaGlu---AspIleValLeuSerAsn 66
Db 409 AACCGGTGATGGAAGACGGGAAATGTACATCGAGAACTAATATACCTAAACACAT 468
Qy 67 MetArgValAlaPheAspArgMetPheLeuAspVallysGluProPheGluPheSerPro 86
Db 469 AAAAAAATA-----GATCGTGTGTTTTA-----GAACCAAGTGATGTTGAACCA 513
Qy 87 TyrHisGluAlaIle-----LeuGluProPheAsnTyrTyrMetPheGlyGlnAsnTyr 104
Db 514 ATGATGAGCGATAGAGCTTTAGACACAGCATTTAATGTCITTAGACACAGGT--- 570
Qy 105 IleArgProLeuValAsnPheArgGluSerTyrValGlyAsnValSerValPheGlyVal 124
Db 571 -----TCATTATATACAAAGTGTATATCAAAATTTATGTCACAAAGGTATT 615
Qy 125 MetGluGlnLeuLysGlnGlyAspLysValValLeuLysSerAsnHisGlnThrGlu 144
Db 616 TCAGAAGCATTTATACGTACATCTGCTCCAAACTTTATGATCTAATGTTATGACACAA 675
Qy 145 AlaAspProAlaValIleAlaLeuMetLeuGluThrThrAsnProHisIleSerGluAsn 164
Db 676 CCAGGC-----GAGACTGATAATTTATGATGTCMAAGACAT 711
Qy 165 IleIleTyrVal-----AlaGlyAspArgValIleThrAspProLeuCysLysPro 181
Db 712 ATTGATGCACCTTACTCGACAAAGTTGGTGACCACTTTATGATTTTGTATATGATGCTCA 771
Qy 182 PheSerMetGlyArgAsnLeuLysValTyrSerLysLysHisMetAsnAspValPro 201
Db 772 GAATCCTATAGTAAGATGTTTTCACAGGATATGAGAAAGAAAGAAATTCGAACCGTAGCA 831
Qy 202 GluLeuAlaGluMetLysLysArgSerAsnThrArgSerLeuLysGluMetAlaLeuLeu 221
Db 832 GTACATAAAGAACAAATTAAGATAGTGGATTTAGAGTTTAAACGGCATCTCAATTAGTT 891
Qy 222 -----LeuArgGlyGlySerLysIleIle 229
Db 892 GAAATATCTAATGAACACTATGTGACAGACATAACACAAAGTATTA 936

RESULT 12
US-09-221-017B-672/c
; Sequence 672, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
```



```

Db 2364 TTCCTTCATCCGTCATTCACCTCCGCAAGGCTTTT-----GATCTTGCCTTCTTT--- 2314
QY 124 ValMetGluGluGlnLeuLysGlnGlyAspLysValVal-----LeuLysSerAsn 140
Db 2313 -----GATCAAGAGGCAAGCGGTGATCTGTGTGTGGAGCGAT 2272
QY 141 HisGlnThrGluAlaAspProAlaValIleAlaLeuMet-----LeuGluThrThr 157
Db 2271 CATGCCACAGATGTCACCCAGTCGCGTCCCTCATTTGTGAGTCGGCTTGAGAGCACT 2212
QY 158 AsnProHisIle-----SerGluAsnIleIleTyrValAlaGlyAspArgValIle 174
Db 2211 CAG-----CACCTCTCATCCGTACGCGCAAGCAACATCTTCAAGCGCTTTTCATGATGAATT 2155
QY 175 ThrAspProLeuCys-----LysProPheSerMetGlyArgAsn 187
Db 2154 GCGGAAACACTGTTTCAGCGCCCATCTCATAGAACATCAGCCCTTGTCCCAAGCGCAGA 2095
QY 188 LeuLeuCysValTyrSerLysLysHis 196
Db 2094 AC-GTATGCCCATATCATGATAGACAT 2069

RESULT 13
US-08-844-054-1
; Sequence 1, Application US/08844054
; Patent No. 6071892
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6071892el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,054
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9608001.5
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31454
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-844-054-1

Alignment Scores:
Pred. No.: 0.0304
Score: 99.00
Percent Similarity: 42.93%
Best Local Similarity: 22.83%
Query Match: 5.08%
Length: 1257
Matches: 42
Conservative: 37
Mismatch: 61
Indels: 45

```

```

DB: 3 Gaps: 10
US-09-926-805-7 (1-374) x US-08-844-054-1 (1-1257)
QY 15 SerGluGlnAspLeuLeuSerGlyIle-----GlnArgGluLeuGluAlaGlyThrLeu 32
Db 665 TCACAGATGCAACTGGTAAATAATTTGCTAAATCAGAAGGAATCCGCTCGCTCAATC 724
QY 33 ProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeu 52
Db 725 CGAAAGAACACTTCTCCATAC-----GAAATGTACCAATTTCTGGATGAAC----- 768
QY 53 GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAlaAlaPheAsp 72
Db 769 -----GTGATGGAGCGCTGACGCTGTTTCGCTTCTTG 798
QY 73 ArgMetPheLeu-----AspValLysGluProPheGluPhe 84
Db 799 AAAATCTTTACTTCTTCTGTCATCTTGATGAGATTGAAGATATTCTGTAACAATTTGAAGCA 858
QY 85 SerProTyr-----HisGluAlaIleLeuGluProPheAsnTyrTyrMetPhe 100
Db 859 GCGCCACACGACGCTTCGCTCAAAAGTCTTGGCTCGTGAAGTTGTACACTTCTTTCAC 918
QY 101 GlyGlnAsnTyrIleArgProLeuValAsnPheArgGluSer---TyrValGlyAsnVal 119
Db 919 GGAGAGAGAGCGCTACAAAGAGCACTTAACATCATCTGAGCAACTCTTTGCGAGGAACATC 978
QY 120 SerValPheGlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuLys 139
Db 979 AAAAACCTTCTGTCAAAGAG-----CTCAACAAGGACTTCGTGTGTG-----CCA 1026
QY 140 AsnHisGlnThrGluAlaAspPro-----AlaValIleAlaLeuMetLeuGluThr--- 156
Db 1027 AACTACCAAGTACAGGCAGAGCAAAACAAATATCGTGGAACCTCTGCTCTCATCGT 1086
QY 157 -----ThrAsnProHisIleSerGluAsnIleIleTyrValAla 169
Db 1087 ATAGTTAACTCAAAACGCCAGCCGCTGAAGACGTCCTCAAAACGGAGCCATCTACGTAAAC 1145
QY 170 GlyAspArgVal 173
Db 1147 GCGGACCGCATC 1158

RESULT 14
US-09-347-333-1
; Sequence 1, Application US/09347333
; Patent No. 6284489
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6284489el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/347,333
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,054
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gimmi, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31454
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1257 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 US-09-347-333-1

Alignment Scores:
 Pred. No.: 0.0304 Length: 1257
 Score: 99.00 Matches: 42
 Percent Similarity: 42.93% Conservative: 37
 Best Local Similarity: 22.83% Mismatches: 61
 Query Match: 5.08% Indels: 45
 DB: 3 Gaps: 10

US-09-926-805-7 (1-374) x US-09-347-333-1 (1-1257)

QY 15 SerGluGlnAspLeuSerGlyLe-----GlnArgGluLeuGluAlaGlyThrLeu 32
 Db 665 TCACAGATGCACTGGTAAATAATTTGGTAAATCAGAAGAAATCGCGTCTGCTCAATC 724
 QY 33 ProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeu 52
 Db 725 CCGAAAGACTTCTCCATAC-----GAAATGTACCAATTCGTGATGAAC----- 768
 QY 53 GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAlaPheAsp 72
 Db 769 -----GTGATGACGCTGACGCTGCTGCTTCTG 798
 QY 73 ArgMetPheLeu-----AspValLysGluProPheGluPhe 84
 Db 799 AAATCTTTACTTTCTTGTCTACTGTGATGAGATTAAGATTCGTAACAACTTTGAAGCA 858
 QY 85 SerProTyr-----HisGluAlaLeuGluProPheAsnTyrTyrMetPhe 100
 Db 859 GCGCCACAGACGCTTGGCTCAAAAGCTTGGCTCGTGAAGTTGTTACACTTGTTCAC 918
 QY 101 GlyGlnAsnTyrIleArgProLeuValAsnPheArgGluSer---TyrValGlyAsnVal 119
 Db 919 GGAGAAGAAGCCCTACAAAGAAGCACCTTAACATCAGTACGACCACTCTTTCGAGAAACATC 978
 QY 120 SerValPheGlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuIleSer 139
 Db 979 AAAACCTTTCTGTCAAGAG-----CTCAACAGAGACTTCGTGCTGTG-----CCA 1026
 QY 140 AsnHisGlnThrGluAlaAspPro-----AlaValIleAlaLeuMetLeuGluThr--- 156
 Db 1027 AACTACCAAGTACAGGACAGCAAGCAACAATATCGTGAATCTGCTGCTCATCTGTGT 1086
 QY 157 -----ThrAsnProHisIleSerGluAsnIleIleTyrValAla 169
 Db 1087 ATAGTTAACTCAAAAGCCGAGCCGCTGAAGACGTCCTCAAAACGAGCCATCTACGTAAC 1146
 QY 170 GlyAspArgVal 173
 Db 1147 GCGACCGCATC 1158

RESULT 15

US-08-961-527-143
 Sequence 143, Application US/08961527
 Patent No. 6420135
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,527
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 143:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4965 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-143

Alignment Scores:
 Pred. No.: 0.401 Length: 4965
 Score: 98.00 Matches: 42
 Percent Similarity: 42.93% Conservative: 37
 Best Local Similarity: 22.83% Mismatches: 61
 Query Match: 5.03% Indels: 45
 DB: 4 Gaps: 10

US-09-926-805-7 (1-374) x US-08-961-527-143 (1-4965)

QY 15 SerGluGlnAspLeuSerGlyLe-----GlnArgGluLeuGluAlaGlyThrLeu 32
 Db 3313 TCACAGATGCACTGGTAAATAATTTGGTAAATCAGAAGAAATCGCGTCTGCTCAATC 3372
 QY 33 ProLysHisIleAlaGlnAlaMetGluGluLeuTyrGlnAsnTyrLysAsnAlaValLeu 52
 Db 3373 CCGAAAGACTTCTCCATAC-----GAAATGTACCAATTCGTGATGAAC----- 3416
 QY 53 GlnSerAlaAlaProHisAlaGluAspIleValLeuSerAsnMetArgValAlaPheAsp 72
 Db 3417 -----GTGATGACGCTGACGCTGCTGCTTCTG 3446
 QY 73 ArgMetPheLeu-----AspValLysGluProPheGluPhe 84
 Db 3447 AAATCTTTACTTTCTTGTCTACTGTGATGAGATTAAGATATTCGTAACAACTTTGAAGCA 3506
 QY 85 SerProTyr-----HisGluAlaLeuGluProPheAsnTyrTyrMetPhe 100
 Db 3507 GCGCCACAGACGCTGGCTCAAAAGTCTTGGCTCGTGAAGTTGTTACACTTGTTCAC 3566
 QY 101 GlyGlnAsnTyrIleArgProLeuValAsnPheArgGluSer---TyrValGlyAsnVal 119
 Db 3567 GGAGAAGAAGCCCTACAAAGAAGCACCTTAACATCAGTACGACCACTCTTTCGAGGAAACATC 3626
 QY 120 SerValPheGlyValMetGluGluGlnLeuLysGlnGlyAspLysValValLeuIleSer 139
 Db 3627 AAAAACCTTTCTGTCAAGAG-----CTCAACAGAGACTTCGTGCTGTG-----CCC 3674

```
QY 140 AsnHisGlnThrGluAlaAspPro-----AlaValIleAlaLeuMetLeuGluThr--- 156
Db |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
3675 AACTACCAAGTACAGGCAGACGAAACAAACAATATCGTGGAACTGCTCGTCTCATCTGGT 3734
QY 157 -----ThrAsnProHisIleSerGluAsnIleIleTyrValAla 169
Db |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
3735 ATAGTTAACTCAAAACGCCAAGCCCGTGAAGACGTCCTCAAAACGGAGGCATCTACGTTAAAC 3794
QY 170 GlyAspArgVal 173
Db |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::||| |||::|||
3795 GCGACCGCATC 3806
```

Search completed: March 16, 2004, 04:44:37
Job time : 608 secs